Regular graphs and the spectra of two-variable logic with counting*

(Extended abstract)

Eryk Kopczyński
University of Warsaw
Email: erykk@mimuw.edu.pl

Tony Tan
Hasselt University and Transnational University of Limburg
Email: ptony.tan@gmail.com

Abstract

The *spectrum* of a first-order logic sentence is the set of natural numbers that are cardinalities of its finite models. This notion has been shown to be closely linked with complexity theory. In fact, asking whether the spectra of first-order logic sentences is closed under complement is equivalent to asking whether NE = CO-NE.

In this paper we show that when restricted to using only two variables, the spectra of first-order logic sentences are closed under complement. In particular, we show that they are semilinear. At the heart of our proof are semilinear characterisations for the existence of regular graphs, the class of graphs in which there are a priori bounds on the degrees of the vertices.

1 Introduction

Descriptive complexity is the study of complexity theory via the methods of mathematical logic. In essence, instead of asking how much time/space needed to check whether the input has a certain property p, in descriptive complexity we ask how hard it is to express property p in a formal language. When we use first-order logic, descriptive complexity captures exactly the important complexity classes. (A beautiful article [20] by Immerman in the Notices of AMS can serve as a gentle introduction to the topic. See also [3, 21, 26] for some well known texts.)

The origin of descriptive complexity can actually be traced back to an innocent question asked by Scholz in [34] when he introduced the notion of the *spectrum* and asked whether there exists a necessary and sufficient condition for a set to be a spectrum. The spectrum of a first-order sentence ϕ, denoted by $\textsc{Spec}(\phi)$, is the set of natural numbers that are cardinalities of finite models of ϕ. Or, more formally, $\textsc{Spec}(\phi) = \{n \mid \text{there is a model of } \phi \text{ of size } n\}$. A set is a *spectrum*, if it is the spectrum of a first-order sentence.

Since its publication, Scholz's question and many of its variants have been investigated by many researchers for the past 60 years. One of the arguably main open problems in this area

*This article has been submitted to arXiv.org under the same title.

is the one asked by Asser in [1], known as *Asser's conjecture*, whether the complement of a spectrum is also a spectrum.

At first glance the notion of spectrum and complexity theory have nothing to do with each other. Hence, it came as a surprise when Jones and Selman [22], as well as Fagin [4] independently showed that a set of integers is a spectrum if and only if its binary representation is in NE. Hence, Asser's conjecture is equivalent to asking whether NE $=$ co-NE. It also immediately implies that if Asser's conjecture is false, i.e., there is a spectrum whose complement is not a spectrum, then NP $\neq$ co-NP, hence P $\neq$ NP. The converse implication is still open. An interesting result in [37] states that if spectra are precisely rudimentary sets, then NE $=$ co-NE and NP $\neq$ co-NP.* There are a number of interesting connections between spectrum and various models of computation such as RAM as well as intrinsic computational behavior. See, for example, [9, 10, 11, 27, 33]. We refer the reader to [2] for a more comprehensive treatment on the spectra problem and its history.

The notion of spectra was then generalised by Fagin [4], called *generalised spectra*, which are the projective classes of Tarski restricted to finite structures. This notion is what later on gives birth to descriptive complexity and finite model theory [20, 5, 3, 21, 26].

The main result in this paper. Since the original Asser's conjecture is a very difficult problem, it is natural then to ask whether it holds on a certain subclass of first-order logic. To this end, we consider the *two-variable logic with counting*, denoted by $\mathsf{FO}^2\mathsf{C}$, which is a class of first-order sentences using only two variables and allowing counting quantifiers $\exists^k z\ \phi(z)$, where $k \geq 1$. Intuitively, $\exists^k z\ \phi(z)$ means that there exist at least k number of z's such that $\phi(z)$ holds. Two-variable logic and its variant is an important class of logic used in many settings in computer science such as verification, specification, artificial intelligence, etc [12, 13, 31, 36]. The class $\mathsf{FO}^2\mathsf{C}$ itself was first introduced and studied in [14, 32].

The following are typical instances of structures expressible in $\mathsf{FO}^2\mathsf{C}$.†

(Ex.1) d-regular graphs: the graphs in which every vertex has degree d.

(Ex.2) (c,d)-biregular graphs: the bipartite graphs on the vertices $U \cup V$, where the degree of each vertex in U and V is c and d, respectively.

(Ex.3) (c,d)-directed-regular graphs: the graphs in which the in-degree and the out-degree of each vertex is c and d, respectively.

An observation from basic graph theory tells us that for "big enough" M and N,‡

(C1) there is a d-regular graph of N vertices if and only if Nd is even number;

(C2) there is a (c,d)-biregular graph in which M vertices are of degree c and N vertices of degree d if and only if $Mc = Nd$;

(C3) there is a (c,d)-directed-regular graph of N vertices if and only if $Nc = Nd$, and hence, $c = d$.

These characterisations immediately imply that the spectra of the sentences (Ex.1)–(Ex.3) above are arithmetic sequences.

*It should be noted that the class of rudimentary sets corresponds precisely to *linear time hierarchy* – the linear time analog of polynomial time hierarchy [38].

† Though the result in this paper holds for arbitrary structures, it helps to assume that the structures of $\mathsf{FO}^2\mathsf{C}$ are graphs in which the vertices and the edges are labelled with a finite number of colours.

‡ "Big enough" means that M and N are greater than a constant K which depends only on c and d.

In this paper we show that the spectra of $\mathsf{FO}^2\mathsf{C}$ are semilinear sets – sets that are finite unions of arithmetic sequences. The main technical difficulty that we encounter here is that we cannot rely on any available "pumping" argument, argument which is commonly used to prove that a set is semilinear. Hence, we proceed by constructing a Presburger formula that captures the spectrum of a given formula. Since Presburger formulas are known to express precisely semilinear sets [8], our result follows immediately.

The crux of our construction is a generalisation of the characterisations (C1)–(C3) above to the following setting. Let $\mathcal{C}$ be a set of ℓ-colors, denoted by $\mathsf{col}_1, \mathsf{col}_2, \ldots, \mathsf{col}_\ell$, and let D be an $\ell \times m$-matrix whose entries are all non-negative integers. We say that a graph $G = (V, E)$ is D-regular, if we can color its edges with colors from $\mathcal{C}$ such that there is a partition $V = V_1 \cup \cdots \cup V_m$, where for every integer $1 \leq i \leq m$, for every vertex $v \in V_i$, for every $1 \leq j \leq \ell$, the number of edges with color col_j adjacent to v is *precisely* $D_{j,i}$. Our setting also allows us to say that the number of edges with color col_j adjacent to v is *at least* $D_{j,i}$. In Theorem 4.1 we obtain a Presburger formula that characterises the set $\{N \mid \text{there is a } D\text{-regular graph of } N \text{ vertices}\}$. In a similar manner, we can define (C, D)-biregular graphs and (C, D)-directed-regular graphs and obtain their respective Presburger formula. [§] We then proceed to observe that the relations in every model of a $\mathsf{FO}^2\mathsf{C}$ formula can be partitioned in such a way that every part forms a (C, D)-directed-regular graph, and every two parts a (C, D)-biregular graph. Applying the Presburger formula that characterises the existence of these regular graphs, we obtain the semilinearity of the spectra of $\mathsf{FO}^2\mathsf{C}$ formulae.

For the converse direction, it is not that difficult to show that every semilinear set is also a spectrum of an $\mathsf{FO}^2\mathsf{C}$ sentence. Since semilinear sets are closed under complement, this establishes the fact that the spectra of $\mathsf{FO}^2\mathsf{C}$ are closed under complement.

It can also be deduced immediately from our proof that the many-sorted spectra of $\mathsf{FO}^2\mathsf{C}$ are also semilinear. We will define precisely many-sorted spectra in Section 2. Intuitively, the *many-sorted* spectra of a formula are spectra which counts the cardinality of the unary predicates in the models of the formula, instead of just counting the sizes of the models.

Moreover, our result extends trivially to the class $\exists\mathsf{SO}^2\mathsf{C}$, the class of sentences of the form:

$$\exists R_1 \cdots \exists R_m \ \phi, \tag{1}$$

where $R_1, \ldots, R_m$ are second-order variables and ϕ uses only two first-order variables. We simply regard $R_1, \ldots, R_m$ as part of the signature. Properties expressible in $\exists\mathsf{SO}^2\mathsf{C}$ include many NP-complete problems such as 3-colorability and k-clique. In fact, in descriptive complexity the class NP is defined precisely by $\exists\mathsf{SO}$ sentences, that is, sentences of the same form (1), but without any restriction on the number of variables in ϕ.

Related works. The logic $\mathsf{FO}^2\mathsf{C}$ is not the first logic known to have semilinear spectra. A well known Parikh theorem states the spectra of context-free languages are semilinear, and closed under complementation. Using the celebrated composition method, Gurevich and Shelah in [16] showed that the spectra of monadic second order logic with *one unary function*

[§] Closely related to our result is the work by S. L. Hakimi [17] which deals with the question: given a vector $(d_1, \ldots, d_m)$, is there a graph with vertices $v_1, \ldots, v_m$ whose degrees are precisely $d_1, \ldots, d_m$, respectively? Another related result concerns the notion of score sequence obtained by H. G. Landau [23] which deals with the question: given a vector $(d_1, \ldots, d_m)$, is there a tournament with vertices $v_1, \ldots, v_m$ whose outdegrees are precisely $d_1, \ldots, d_m$, respectively? These questions are evidently different from our characterisations provided in Sections 4, 5 and 7.

are semilinear. In [7] Fischer and Makowsky show that the spectra of the monadic second-order logic with *modulo counting* over structures with *bounded tree-width* are semilinear.

On the other hand, structures expressible in $\mathsf{FO}^2\mathsf{C}$ do not have bounded tree-width. An example is d-regular graphs for $d \geq 3$. It should also be noted that in $\mathsf{FO}^2\mathsf{C}$ one can express a few unary functions, hence our result does not follow from [16], and neither theirs from ours since we are restricted to using only two variables. A significant difference between our result and the previous ones is we do not put any restriction on the vocabulary nor in the interpretation.

The result closest to ours is the one by É. Grandjean in [11] where he considers the spectra of first-order sentences using only one variable. A similar result due to M. Grohe and stated in [2], says that for every Turing machine M, there exists a first-order sentence ϕ_M using only *three variables* such that $\textsc{Spec}(\phi_M) = \{t^2 \mid t \text{ is the length of an accepting run of } M\}$. With this evidence it is widely believed that every spectrum is the spectrum of a first-order sentence using only three variables [2], thus, making our result "almost optimal."

It is a standard practice in logic to relate the number of variables with the intrinsic computational behaviour. See, for example, [19, 31]. It was known that two-variable logic has small model property and hence decidable [30]. On the other hand, $\mathsf{FO}^2\mathsf{C}$ does not have finite model property, but still decidable [14, 32]. In the context of specification and verification of concurrent system, two-variable logic (even without counting) is especially an important class [13] due to its relation to modal logic commonly used in specification and artificial intelligence. Our result and its proof give another interesting insight why modal logic is so robustly decidable [12, 36] – each of its model is simply a collection of regular graphs. There is also a related result in [6, 35] where it was shown how to encode in polynomial time the derivation trees of a context-free grammar to Presburger formula.

Outline of the paper. This paper is organised as follows. In Section 2 we review the definition of $\mathsf{FO}^2\mathsf{C}$ and state our main results and a few corolarries. In Section 3 we present the modal logic for $\mathsf{FO}^2\mathsf{C}$, called *quantified modal logic with counting* (QMLC). For our purpose it is easier to work with QMLC than with $\mathsf{FO}^2\mathsf{C}$. In Sections 4 and 5 we introduce the notion of regularity and biregularity for *undirected* graphs and present their Presburger characterisations, which will then be used in Section 6 to show the semilinearity of the spectra of $\mathsf{FO}^2\mathsf{C}$ when restricted to undirected graphs. In Section 7 we introduce the notion of regularity for *directed* graphs and present their Presburger characterisations, which will then be used in Section 8 to prove the semilinearity of the spectra of the full $\mathsf{FO}^2\mathsf{C}$. Obviously, the result in Section 6 is subsumed in Section 8. However, we believe that it is interesting enough to be stated independently. Finally we conclude with some concluding remarks in Section 9. The details of most proofs are put in the appendix.

2 The logic $\mathsf{FO}^2\mathsf{C}$

In this section we review the definition of $\mathsf{FO}^2\mathsf{C}$ and mention the main result in this paper and its corollaries. In the following let $\mathcal{P} = \{P_1, P_2, \ldots\}$ be the set of predicate symbols of arity 1; while $\mathcal{R} = \{R_1, R_2, \ldots\}$ the set of predicate symbols of arity 2. Two-variable logic with counting ($\mathsf{FO}^2\mathsf{C}$) is defined by the following syntax.

$$\phi \quad := \quad \neg\phi \;\big|\; R(z,z) \;\big|\; P(z) \;\big|\; \phi_1 \wedge \phi_2 \;\big|\; \exists^k z\ \phi$$

where the variable z ranges over x, y, R over $\mathcal{R}$ and P over $\mathcal{P}$.

The quantifier $\exists^k z\ \phi$ means that *there are at least k elements z* such that ϕ holds. Note that $\exists^1 z\ \phi$ is equivalent to the standard $\exists z\ \phi$, and $\forall z\ \phi$ is equivalent to $\neg\exists^1 z\ \neg\phi$. By default, we assume that $\exists^0 z\ \phi$ always holds.

For simplicity, we are going to use the following notations.

$$\begin{aligned}\exists^{=k} z\ \phi &:= \exists^k z\ \phi \ \wedge\ \neg(\exists^{k+1} z\ \phi)\\ \exists^{\leq k} z\ \phi &:= \neg(\exists^{k+1} z\ \phi)\end{aligned}$$

As usual, we write $\mathfrak{A} \models \phi$, if the structure $\mathfrak{A}$ is a model of ϕ and $\mathsf{Model}(\phi) = \{\mathfrak{A} \mid \mathfrak{A} \models \phi\}$.

Theorem 2.1 below is the main result in this paper. We present its proof is Section 8.

Theorem 2.1 *For every $\phi \in \mathsf{FO}^2\mathsf{C}$, there exists a Presburger formula $\Psi(x)$ such that the set $\{n \mid \Psi(n)$ holds$\} = \textsc{Spec}(\phi)$.*

We should remark that Theorem 2.1 also holds for arbitrary vocabulary. Since $\mathsf{FO}^2\mathsf{C}$ uses only two variables, relations of greater arity such as $R(x, y, x, x, y)$ can be viewed simply as unary or binary relations; so we can create new binary and unary relations for each possible combination, and easily verify whether the result is consistent.

An immediate consequence of Theorem 2.1 is the spectra of $\mathsf{FO}^2\mathsf{C}(\mathsf{Sym})$ are semilinear, since it is known from [8] that every set expressible by Presburger formula is semilinear.

Corollary 2.2 *For every sentence $\phi \in \mathsf{FO}^2\mathsf{C}$, the spectrum $\textsc{Spec}(\phi)$ is semilinear.*

On the other hand, it is not that difficult to show that every semilinear set is a spectrum of an $\mathsf{FO}^2\mathsf{C}$ sentence, as formally stated below.

Proposition 2.3 *For every semilinear set $\Lambda \subseteq \mathbb{N}$, there exists a sentence $\phi \in \mathsf{FO}^2\mathsf{C}$ such that $\textsc{Spec}(\phi) = \Lambda$.*

Combining Corollary 2.2 and Proposition 2.3, we obtain the following corollary.

Corollary 2.4 *The spectra of $\mathsf{FO}^2\mathsf{C}$ sentences are closed under complement within $\mathsf{FO}^2\mathsf{C}$.*

Proof. It follows from Corollary 2.2 and Proposition 2.3 and the fact that semilinear sets are closed under complement. ■

Theorem 2.1 can be further generalised as follows. We define the *many-sorted spectrum* as the Parikh image of the graph G, where $\mathsf{Parikh}(G) = (n_1, \ldots, n_l)$, and n_i is the number of nodes in G labeled with P_i. With a slight adjustment in our proof in Section 8, we can obtain the following corollary.

Corollary 2.5

- *The Parikh image of the graphs satisfying a formula $\phi \in \mathsf{FO}^2\mathsf{C}$ is semilinear.*
- *The following problem is decidable. Given an $\mathsf{FO}^2\mathsf{C}$ formula ϕ with $\mathcal{P} = \{P_1, \ldots, P_l\}$ be the set of unary predicates and a Presburger formula $\Psi(x_1, \ldots, x_l)$, determine whether there exists a graph $G \in \mathsf{Model}(\phi)$ such that $\Psi(\mathsf{Parikh}(G))$ holds.*

3 Quantified modal logic with counting

In this section we present an equivalent form of $\mathsf{FO}^2\mathsf{C}$, called *quantified modal logic with counting* (QMLC). For our purpose, it will be easier to work on QMLC rather than $\mathsf{FO}^2\mathsf{C}$ directly.

Moreover, for ease of presentation, we make the following assumption. Let $\phi \in \mathsf{FO}^2\mathsf{C}$ and let $\mathcal{R}$ and $\mathcal{S}$ be the set of binary and unary predicates appearing in ϕ, respectively. By extending $\mathcal{R}$ and $\mathcal{S}$ and by modifying the sentence ϕ, if necessary, we can construct another $\phi' \in \mathsf{FO}^2\mathsf{C}$ such that the structures of ϕ' have the same sizes as those of ϕ and every structure $\mathfrak{A} \in \mathsf{Model}(\phi')$ satisfies the following.

- $\mathfrak{A}$ is a clique over A.
  That is, for every $a, b \in A$, either $a = b$ or $R(a, b)$ for some $R \in \mathcal{R}$.
- Every binary relation in $\mathcal{R}$ is not reflexive.
  That is, for every $R \in \mathcal{R}$, for every $a, b \in A$, if $R(a, b)$, then $a \neq b$.
- $\mathcal{R}$ is closed under reversal.
  That is, for every $R \in \mathcal{R}$, there exists $\overleftarrow{R} \in \mathcal{R}$ such that $\overleftarrow{R} \neq R$ and for every $a, b \in A$, $R(a, b)$ if and only if $\overleftarrow{R}(b, a)$.
- The binary predicates in $\mathcal{R}$ are pairwise disjoint.

Note that all these assumptions can be written in $\mathsf{FO}^2\mathsf{C}$ sentences.

The class MLC of *modal logic with counting* is defined with the following syntax.

$$\phi \quad ::= \quad \neg\phi \;\big|\; \alpha \;\big|\; \phi_1 \wedge \phi_2 \;\big|\; \Diamond_\beta^k \phi$$

where α ranges over Σ and β over $\mathcal{R}$.

The semantics of MLC is as follows. Let $\mathfrak{A}$ be a structure of τ and $a \in A$ and ϕ be an MLC formula. That $\mathfrak{A}$ satisfies ϕ from a, denoted by $\mathfrak{A}, a \models \phi$, is defined as followed.

- $\mathfrak{A}, a \models P$, where $P \in \mathcal{P}$, if $S(a)$ holds in $\mathfrak{A}$.
- $\mathfrak{A}, a \models \neg\phi$, if $\mathfrak{A}, a \not\models \phi$.
- $\mathfrak{A}, a \models \phi_1 \wedge \phi_2$, if $\mathfrak{A}, a \models \phi_1$ and $\mathfrak{A}, a \models \phi_2$.
- $\mathfrak{A}, a \models \Diamond_R^k \phi$, if there exist at least k elements $b_1, \ldots, b_k \in A$ such that $R(a, b_i)$ holds in $\mathfrak{A}$ and $\mathfrak{A}, b_i \models \phi$ for $i = 1, \ldots, k$.

We define the class of quantified modal logic with counting, denoted by QMLC with the following syntax.

$$\phi \quad ::= \quad \neg\phi \;\big|\; \phi_1 \wedge \phi_2 \;\big|\; \exists^k \psi$$

where the formula $\psi \in \mathsf{MLC}$. A QMLC formula ϕ is called a *basic* QMLC, if it is of the form $\exists^k\ \varphi$, where $\varphi \in \mathsf{MLC}$.

The semantics of QMLC is as follows. Let $\mathfrak{A}$ be a structure of τ and $\phi \in \mathsf{QMLC}$. That $\mathfrak{A}$ satisfies ϕ, denoted by $\mathfrak{A} \models \phi$, is defined as followed.

- $\mathfrak{A} \models \neg\phi$, if it is not the case that $\mathfrak{A} \models \phi$.
- $\mathfrak{A} \models \phi_1 \wedge \phi_2$, if $\mathfrak{A} \models \phi_1$ and $\mathfrak{A} \models \phi_2$.
- $\mathfrak{A} \models \exists^k \psi$, if there exist at least k elements $a_1, \ldots, a_k \in A$ such that $\mathfrak{A}, a_i \models \psi$ for $i = 1, \ldots, k$.

We also denote by $\textsc{Spec}(\phi)$, the spectrum of a QMLC formula ϕ. It can be shown by a technique similar to the one in [28] that $\mathsf{FO}^2\mathsf{C}$ and QMLC are equivalent. (See Appendix A for the details.)

The following notion of *type* will be useful. Let ϕ be a $\mathsf{QMLC}(\mathsf{Sym})$ sentence. Let $\mathcal{M}_\phi$ be the set of all MLC subformulae of ϕ and their negations. A *type* in ϕ is a subset $T \subseteq \mathcal{M}_\phi$ such that

- if $\varphi_1 \wedge \varphi_2 \in T$, then both $\varphi_1, \varphi_2 \in T$;
- $\varphi \in T$ if and only if $\neg\varphi \notin T$;
- if $\neg(\varphi_1 \wedge \varphi_2) \in T$, then at least one of $\neg\varphi_1, \neg\varphi_2 \in T$.

For a structure $\mathfrak{A}$ and an element $a \in A$, we define *the type of a in $\mathfrak{A}$*, denoted by $\mathsf{type}_{\mathfrak{A}}(a) \subseteq \mathcal{M}_\phi$, where $\varphi \in \mathsf{type}_{\mathfrak{A}}(a)$ if and only if $\mathfrak{A}, a \models \varphi$. For a type T, we write $T(\mathfrak{A})$ to denote the set of elements in A with type T. Note that the sets $T(\mathfrak{A})$'s are disjoint. We let $\mathcal{T}_\phi$ to be the set of all types in ϕ.

4 Undirected regular graphs

In this section and the next we introduce the notion of regularity and biregularity for undirected graphs. It should be noted that both classes of graphs are expressible in $\mathsf{FO}^2\mathsf{C}$ sentences. In the following $\mathbb{N}$ denote the set $\{0, 1, 2, \ldots\}$.

We start with the notion of ℓ-type graphs. An ℓ-type undirected graph is a tuple $G = (V, E_1, \ldots, E_\ell)$, where $E_1, \ldots, E_\ell$ are *pairwise disjoint* symmetric relations on V. Edges in E_i are called E_i-edges. It helps to think of an ℓ-type graph as a graph in which the edges are coloured with ℓ number of colours. We write $V(G)$ to denote the set of vertices V in G and $E_i(G)$ the set E_i, for each $i = 1, \ldots, \ell$.

For a vertex $u \in V(G)$, $\deg_{E_i}(u)$ denotes the number of E_i-edges adjacent to it, and $\deg(u) = \sum_{i=1}^{\ell} \deg_{E_i}(u)$. We write $\deg(G) = \max\{\deg(u) \mid u \text{ is a vertex in } G\}$. For an integer $d \in \mathbb{N}$, we write $\deg_{E_i}(u) = {}^{\blacktriangleright}d$, if $\deg_{E_i}(u) \geq d$.

Let ${}^{\blacktriangleright}\mathbb{N} = \{{}^{\blacktriangleright}1, {}^{\blacktriangleright}2, \ldots\}$ and $\mathbb{B} = \mathbb{N} \cup {}^{\blacktriangleright}\mathbb{N}$. We write $\mathbb{B}^{\ell \times m}$ to denote the set of $\ell \times m$ matrices whose entries are from $\mathbb{B}$. The entry in row i and column j of a matrix $D \in \mathbb{B}^{\ell \times m}$ is denoted by $D_{i,j}$. An ℓ-type graph $G = (V, E_1, \ldots, E_\ell)$ is D-regular, if there exists a partition $V = V_1 \cup \cdots \cup V_m$ such that for each $i = 1, \ldots, \ell$, for each $j = 1, \ldots, m$, for each vertex $v \in V_j$, $\deg_{E_i}(v) = D_{i,j}$.[¶] We say that $V_1 \cup \cdots \cup V_m$ is a partition of D-regularity and the graph G is of size $\bar{N} = (N_1, \ldots, N_m)$, if $(N_1, \ldots, N_m) = (|V_1|, \ldots, |V_m|)$.

Theorem 4.1 *For every matrix $D \in \mathbb{B}^{\ell \times m}$, there is a Presburger formula $\mathsf{REG}_D(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. There exists a D-regular ℓ-type graph of size $\bar{N}$ if and only if $\mathsf{REG}_D(\bar{N})$ holds.*

This theorem is then generalised to the case of complete graphs. An ℓ-type graph is a *complete* graph, if every two of its vertices are connected by an edge. In other words, if $G = (V, E_1, \ldots, E_\ell)$ is a *complete* graph, then the graph $G' = (V, E_1 \cup \cdots \cup E_\ell)$ is a clique. If G is also a D-regular graph, then we call G a D-regular-complete graph.

The following theorem is the main result in this section that will be used in Section 6.

[¶]If $D_{i,j} = {}^{\blacktriangleright}d$, then this means $\deg_{E_i}(v) \geq d$.

Theorem 4.2 *For every matrix $D \in \mathbb{B}^{\ell \times m}$, there is a Presburger formula $\textsf{REG-COMP}_D(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. There exists a D-regular-complete graph of size $\bar{N}$ if and only if $\textsf{REG-COMP}_D(\bar{N})$ holds.*

5 Undirected biregular graphs

In this section we introduce the notion of biregularity for undirected bipartite graphs. An ℓ-type bipartite graph is a tuple $G = (U, V, E_1, \ldots, E_\ell)$, where $E_i \subseteq U \times V$, for each $i = 1, \ldots, \ell$.

Let $C \in \mathbb{B}^{\ell \times m}$ and $D \in \mathbb{B}^{\ell \times n}$. An ℓ-type bipartite graph $G = (U, V, E_1, \ldots, E_\ell)$ is (C, D)-biregular, if there is a partition $U = U_1 \cup \cdots \cup U_m$ and $V = V_1 \cup \cdots \cup V_n$ such that the following holds.

- For every $i = 1, \ldots, \ell$, for each $j = 1, \ldots, m$, for each vertex $u \in U_j$, $\deg_{E_i}(u) = C_{i,j}$.
- For every $i = 1, \ldots, \ell$, for each $j = 1, \ldots, n$, for each vertex $v \in V_j$, $\deg_{E_i}(v) = D_{i,j}$.

The partitions $U = U_1 \cup \cdots \cup U_m$ and $V = V_1 \cup \cdots \cup V_n$ are called the partitions of the (C, D)-biregularity of G. We say that the (C, D)-biregular graph G is of size $(\bar{M}, \bar{N})$, if $\bar{M} = (|U_1|, \ldots, |U_m|)$ and $\bar{N} = (|V_1|, \ldots, |V_n|)$.

Theorem 5.1 *For every two matrices $C \in \mathbb{B}^{\ell \times m}$ and $D \in \mathbb{B}^{\ell \times n}$, there is a Presburger formula $\textsf{BiREG}_{C,D}(\bar{X}, \bar{Y})$, where $\bar{X} = (X_1, \ldots, X_m)$ and $\bar{Y} = (Y_1, \ldots, Y_n)$ such that the following holds. There exists an ℓ-type (C, D)-biregular graph of size $(\bar{M}, \bar{N})$ if and only if $\textsf{BiREG}_{C,D}(\bar{M}, \bar{N})$ holds.*

This theorem is then generalised to the case of complete bipartite graphs. An ℓ-type bipartite graph $G = (U, V, E_1, \ldots, E_\ell)$ is *complete*, if the graph $G' = (U, V, E_1 \cup \cdots \cup E_\ell)$ is a complete bipartite graph. If G is also a (C, D)-biregular graph, then we call it a (C, D)-biregular-complete graph.

The following theorem is the main result in this section that will be used in Section 6.

Theorem 5.2 *For every two matrices $C \in \mathbb{B}^{\ell \times m}$ and $D \in \mathbb{B}^{\ell \times n}$, there is a Presburger formula $\textsf{BiREG-COMP}_{C,D}(\bar{X}, \bar{Y})$, where $\bar{X} = (X_1, \ldots, X_m)$ and $\bar{Y} = (Y_1, \ldots, Y_n)$ such that the following holds. There exists a (C, D)-biregular-complete graph of size $(\bar{M}, \bar{N})$ if and only if $\textsf{BiREG-COMP}_{C,D}(\bar{M}, \bar{N})$ holds.*

6 Proof of Theorem 2.1 on undirected graphs

In this section we are going to prove Theorem 2.1 when the interpretations are restricted on undirected graphs. We denote by $\textsf{FO}^2\textsf{C}(\textsf{Sym})$ the subclass of $\textsf{FO}^2\textsf{C}$ formula when all the binary relations are assumed to be symmetric. We let $\textsf{QMLC}(\textsf{Sym})$ to be its equivalent $\textsf{QMLC}$ counterpart.

Obviously the semilinearity of the spectra of $\textsf{FO}^2\textsf{C}(\textsf{Sym})$ is only a special case of Theorem 2.1. Nevertheless, we believe that it is sufficiently interesting to be stated on its own especially since it uses a separate tool developed in Section 4. and the Presburger formulae constructed for both $\textsf{FO}^2\textsf{C}(\textsf{Sym})$ and $\textsf{FO}^2\textsf{C}$ are slightly different. Not surprisingly the formula for $\textsf{FO}^2\textsf{C}(\textsf{Sym})$ is simpler since it can "disregard" the orientation of the edges.

In the remaining of this subsection we are going to present the construction of the desired Presburger formula for $\textsf{QMLC}(\textsf{Sym})$ formula ϕ. We start with basic $\textsf{QMLC}(\textsf{Sym})$ formulae. Recall that a $\textsf{QMLC}$ formula ϕ is a *basic* $\textsf{QMLC}$, if it is of the form $\exists^k\ \varphi$, where $\varphi \in \textsf{MLC}$.

Proposition 6.1 *For every basic* QMLC(Sym) *formula* ϕ*, there exists a Presburger formula* $\mathit{PREB}_\phi(x)$ *such that* $\text{SPEC}(\phi) = \{n \mid \mathit{PREB}_\phi(n) \text{ holds}\}$.

Proof. Let ϕ be of the form: $\exists^k\ \varphi$. Let $\mathcal{R} = \{R_1, \ldots, R_\ell\}$ to be the set of binary relations used in ϕ. Let K be the integer such that for every subformula $\Diamond^l_R\psi$ in ϕ, we have $l \leq K$.

Recall the notion of type introduced in Section 3 and that $\mathcal{T}_\phi$ is the set of all types in ϕ.

A function $f\ :\ \mathcal{T}_\phi \times \mathcal{R} \times \mathcal{T}_\phi \to \{0, 1, \ldots, K\}\ \cup\ \{\blacktriangleright K\}$ is *consistent*, if for every $T \in \mathcal{T}_\phi$ the following holds.

- If $\Diamond^l_R\ \mu \in T$, then $\sum_{S \text{ s.t. } S \ni \mu}\ f(T, R, S) \geq l$.
- If $\neg(\Diamond^l_R\ \mu) \in T$, then $\sum_{S \text{ s.t. } S \ni \mu}\ f(T, R, S) \leq l - 1$.

Intuitively a consistent function is to indicate that the semantics of each type is obeyed. More precisely, it is intended to mean that for an element a of type T there are $f(T, R, S)$ number of R-edges going from the element a to elements of type S. Let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be the set of all consistent functions.

For a type $T \in \mathcal{T}_\phi$, we define the matrix $D_T \in \mathbb{B}^{\ell \times m}$ as follows.

$$D_T \quad := \quad \begin{pmatrix} f_1(T, R_1, T) & f_2(T, R_1, T) & \cdots & f_m(T, R_1, T) \\ f_1(T, R_2, T) & f_2(T, R_2, T) & \cdots & f_m(T, R_2, T) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(T, R_\ell, T) & f_2(T, R_\ell, T) & \cdots & f_m(T, R_\ell, T) \end{pmatrix}$$

For two different types $S, T \in \mathcal{T}_\phi$, we define the matrix $D_{S\to T} \in \mathbb{B}^{\ell \times m}$ as follows.

$$D_{S\to T} \quad := \quad \begin{pmatrix} f_1(S, R_1, T) & f_2(S, R_1, T) & \cdots & f_m(S, R_1, T) \\ f_1(S, R_2, T) & f_2(S, R_2, T) & \cdots & f_m(S, R_2, T) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(S, R_\ell, T) & f_2(S, R_\ell, T) & \cdots & f_m(S, R_\ell, T) \end{pmatrix}$$

Now we define the formula $\mathsf{PREB}_\phi(x)$. First, we enumerate the set $\mathcal{T}_\phi = \{T_1, \ldots, T_n\}$ and the set $\mathcal{T}_\phi \times \mathcal{F} = \{(T_1, f_1), \ldots, (T_n, f_m)\}$. The formula $\mathsf{PREB}_\phi(x)$ is defined as follows.

$$\exists X_{(T_1,f_1)} \cdots \exists X_{(T_n,f_m)} \quad \left(x = \sum_{1\leq i\leq n}\ \sum_{1\leq j\leq m} X_{T_i,f_j} \right)\ \wedge\ \mathsf{PREB\text{-}Atom}_\phi(\bar{X})\ \wedge\ \mathsf{CON}(\bar{X})$$

where $\bar{X} = (X_{(T_1,f_1)}, \ldots, X_{(T_n,f_m)})$ the vector of all the variables $X_{(T,f)}$'s and

$$\begin{aligned} \mathsf{PREB\text{-}Atom}_\phi(\bar{X}) \quad &:= \quad \sum_{(T,f) \text{ s.t. } \varphi\in T} X_{(T,f)} \geq k \\ \mathsf{CON}(\bar{X}) \quad &:= \quad \bigwedge_{T\in\mathcal{T}} \mathsf{REG\text{-}COMP}_{D_T}(\bar{X}_T)\ \wedge \bigwedge_{S,T\in\mathcal{T}} \mathsf{BiREG\text{-}COMP}_{D_{S\to T},D_{T\to S}}(\bar{X}_S, \bar{X}_T) \end{aligned}$$

where $\bar{X}_T = (X_{(T,f_1)}, \ldots, X_{(T,f_m)})$ and $\bar{X}_S = (X_{(S,f_1)}, \ldots, X_{(S,f_m)})$ the vector of variables associated with the types T and S, respectively.

We claim that PREB_ϕ defines precisely the spectrum of ϕ. Intuitively, the meaning of the variable $X_{T,f}$ is the number of elements of type T in which there exists $f(T, R, S)$ number of R-edges going to elements of type S. The formula CON is to make sure that the values for

$X_{T,f}$'s are consistent, that is, they obey the intended meaning of T and f; while $\mathsf{PREB\text{-}Atom}_\phi$ is to make sure that the original QMLC formula ϕ is satisfied. The details can be found in Appendix G. ∎

Proposition 6.2 *Let $\phi \in \mathsf{QMLC(Sym)}$ be a negation of a basic* QMLC *formula. Then there exists a Presburger formula $\mathsf{PREB}_\phi(x)$ such that $\textsc{Spec}(\phi) = \{n \mid \mathsf{PREB}_\phi(n) \text{ holds}\}$.*

Proof. The proof is almost the same as in the proof of Proposition 6.1. Let ϕ be $\neg\exists^k\varphi$, where $\varphi \in \mathsf{MLC}$. We simply replace the formula PREB-Atom with $\sum_{(T,f) \text{ s.t. } \varphi \in T} X_{(T,f)} \leq k-1$. This completes our proof of Proposition 6.2. ∎

Theorem 6.3 *For every $\phi \in \mathsf{FO^2C(Sym)}$, there exists a Presburger formula $\mathsf{PREB}_\phi(x)$ such that the set $\{n \mid \Psi(n) \text{ holds}\} = \textsc{Spec}(\phi)$. In particular, the spectra of $\mathsf{FO^2C(Sym)}$ sentences are closed under complement within $\mathsf{FO^2C(Sym)}$ itself.*

Proof. Let $\phi \in \mathsf{FO^2C(Sym)}$. We may assume that ϕ is a $\mathsf{QMLC(Sym)}$ formula. Now the proof of Theorem 6.3 is a straightforward generalisation of Propositions 6.1 and 6.2. We simply redefine $\mathsf{PREB\text{-}Atom}_\phi$ as follows. First, we push all the negations inside that they are only applied to basic $\mathsf{QMLC(Sym)}$.

We redefine $\mathsf{PREB\text{-}Atom}_\phi$ inductively as follows.

- If $\phi := \exists^k\varphi$, then $\mathsf{PREB\text{-}Atom}_\phi := \sum_{(T,f) \text{ s.t. } \varphi \in T} X_{(T,f)} \geq k$.
- If $\phi := \neg\exists^k\varphi$, then $\mathsf{PREB\text{-}Atom}_\phi := \sum_{(T,f) \text{ s.t. } \varphi \in T} X_{(T,f)} \leq k-1$.
- If $\phi := \phi_1 \wedge \phi_2$, then $\mathsf{PREB\text{-}Atom}_\phi := \mathsf{PREB\text{-}Atom}_{\phi_1} \wedge \mathsf{PREB\text{-}Atom}_{\phi_2}$.
- If $\phi := \phi_1 \vee \phi_2$, then $\mathsf{PREB\text{-}Atom}_\phi := \mathsf{PREB\text{-}Atom}_{\phi_1} \vee \mathsf{PREB\text{-}Atom}_{\phi_2}$.

Now the formula $\mathsf{PREB}_\phi(x)$ is defined exactly like above.

$$\exists X_{(T_1,f_1)} \cdots \exists X_{(T_n,f_m)} \quad \left(x = \sum_{1 \leq i \leq n} \quad \sum_{1 \leq j \leq m} X_{T_i,f_j} \right) \wedge \ \mathsf{PREB\text{-}Atom}_\phi(\bar{X}) \ \wedge \ \mathsf{CON}(\bar{X})$$

where the formula CON is as in Proposition 6.1. The correctness of the formula $\mathsf{PREB}_\phi(x)$ can be verified by a straightforward induction.

On the other hand, it is not that difficult to show that every semilinear set is a spectrum of an $\mathsf{FO^2C(Sym)}$ sentence, hence the spectra of $\mathsf{FO^2C(Sym)}$ is closed under complement within $\mathsf{FO^2C(Sym)}$ itself. This completes our proof of Theorem 6.3. ∎

7 Directed regular graphs

In this section we introduce the notion of regularity for directed graphs.

An ℓ-type directed graph is a tuple $G = (V, E_1, \ldots, E_\ell)$, where $E_1, \ldots, E_\ell$ are *pairwise disjoint* irreflexive relations re on V and for every $u, v \in V$, if $(u,v) \in E_1 \cup \cdots \cup E_\ell$, then the reverse direction $(v,u) \notin E_1 \cup \cdots \cup E_\ell$. Edges in E_i are called E_i-edges.

We will write $\text{in-deg}_{E_i}(u)$ to denote the number of incoming E_i-edges toward the vertex u, and $\text{out-deg}_{E_i}(u)$ to denote the number of outgoing E_i-edges from the vertex u. As before,

for an integer $d \in \mathbb{N}$, we write in-deg$_{E_i}(u) = {}^{\blacktriangleright}d$ and out-deg$_{E_i}(u) = {}^{\blacktriangleright}d$, to indicate that in-deg$_{E_i}(u) \geq d$ and in-deg$_{E_i}(u) \geq d$, respectively.

Let ${}^{\blacktriangleright}\mathbb{N} = \{{}^{\blacktriangleright}1, {}^{\blacktriangleright}2, \ldots\}$ and $\mathbb{B} = \mathbb{N} \cup {}^{\blacktriangleright}\mathbb{N}$. We write $\mathbb{B}^{\ell \times m}$ to denote the set of $\ell \times m$ matrices whose entries are from $\mathbb{B}$. The entry in row i and column j of a matrix $D \in \mathbb{B}^{\ell \times m}$ is denoted by $D_{i,j}$.

Let $C, D \in \mathbb{B}^{\ell \times m}$. An ℓ-type directed graph $G = (V, E_1, \ldots, E_\ell)$ is (C, D)-directed-regular, if there exists a partition $V = V_1 \cup \cdots \cup V_m$ such that for each $i = 1, \ldots, \ell$, for each $j = 1, \ldots, m$, for each vertex $v \in V_j$, in-deg$_{E_i}(v) = C_{i,j}$ and out-deg$_{E_i}(v) = D_{i,j}$.

We say that $V_1 \cup \cdots \cup V_m$ is a partition of (C, D)-regularity and the graph G is of size $\bar{N} = (N_1, \ldots, N_m)$, if $(N_1, \ldots, N_m) = (|V_1|, \ldots, |V_m|)$.

Theorem 7.1 *For every $C, D \in \mathbb{B}^{\ell \times m}$, there exists a Presburger formula $\textsf{Dir-REG}_{C,D}(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. There exists a (C, D)-directed-regular ℓ-type graph of size $\bar{N}$ if and only if $\textsf{Dir-REG}_{C,D}(\bar{N})$ holds.*

An ℓ-type graph $G = (V, E_1, \ldots, E_\ell)$ is a *complete* directed graph, if every two vertices u, v, either (u, v) or (v, u) is in $E_1 \cup \cdots \cup E_\ell$. By the standard graph theoretic term, a *complete* directed graph is a tournament with coloured edges. If G is also a D-regular graph, then we call G a D-directed-regular-complete graph.

Theorem 7.2 *For every $C, D \in \mathbb{B}^{\ell \times m}$, there is a Presburger formula $\textsf{Dir-REG-COMP}_{C,D}(\bar{X})$ such that the following holds. There exists a D-directed-regular-complete graph of size $\bar{N}$ if and only if $\textsf{Dir-REG-COMP}_{C,D}(\bar{N})$ holds.*

8 Proof of Theorem 2.1

In this section we are going to prove Theorem 2.1 using the tools developed in Sections 5 and 7. Recall that a $\textsf{QMLC}$ formula ϕ is a *basic* $\textsf{QMLC}$, if it is of the form $\exists^k \varphi$, where $\varphi \in \textsf{MLC}$.

Proposition 8.1 *For every basic* $\textsf{QMLC}$ *formula ϕ, there exists a Presburger formula $\textsf{PREB}_\phi(x)$ such that $\textsc{Spec}(\phi) = \{n \mid \textsf{PREB}_\phi(n) \text{ holds}\}$.*

Proof. Let ϕ be of the form: $\exists^k \varphi$. Let $\mathcal{R} = \{R_1, \ldots, R_\ell, \overleftarrow{R}_1, \ldots, \overleftarrow{R}_\ell\}$ be the set of binary relations used in ϕ and that $\overleftarrow{R}_i$ is the reversed relation of R_i. Let K be the integer such that for all subformula $\Diamond_R^l \psi$ in ϕ, we have $l \leq K$.

Recall the notion of type introduced in Section 3 and that $\mathcal{T}_\phi$ is the set of all types in ϕ. Recall also the notion of consistent function $f : \mathcal{T}_\phi \times \mathcal{R} \times \mathcal{T}_\phi \to \{0, 1, \ldots, K\} \cup \{{}^{\blacktriangleright}K\}$ introduced in the proof of Proposition 6.1, which for convenience, we repeat here. A function f is *consistent*, if for every $T \in \mathcal{T}_\phi$ the following holds.

- If $\Diamond_R^l\ \mu \in T$, then $\sum_{T' \text{ s.t. } T' \ni \mu} f(T, R, T') \geq l$.
- If $\neg(\Diamond_R^l\ \mu) \in T$, then $\sum_{T' \text{ s.t. } T' \ni \mu} f(T, R, T') \leq l - 1$, and $f(T, R, T') \in \mathbb{N}$, for every $R \in \mathcal{R}$ and for every type $T' \ni \mu$.

Let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be the set of all consistent functions.

For a type $T \in \mathcal{T}_\phi$, we define two matrices $D_T, \overleftarrow{D}_T \in \mathbb{B}^{\ell \times m}$ as follows.

$$D_T \quad := \quad \begin{pmatrix} f_1(T, R_1, T) & f_2(T, R_1, T) & \cdots & f_m(T, R_1, T) \\ f_1(T, R_2, T) & f_2(T, R_2, T) & \cdots & f_m(T, R_2, T) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(T, R_\ell, T) & f_2(T, R_\ell, T) & \cdots & f_m(T, R_\ell, T) \end{pmatrix}$$

and

$$\overleftarrow{D}_T \ := \ \begin{pmatrix} f_1(T,\overleftarrow{R}_1,T) & f_2(T,\overleftarrow{R}_1,T) & \cdots & f_m(T,\overleftarrow{R}_1,T) \\ f_1(T,\overleftarrow{R}_2,T) & f_2(T,\overleftarrow{R}_2,T) & \cdots & f_m(T,\overleftarrow{R}_2,T) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(T,\overleftarrow{R}_\ell,T) & f_2(T,\overleftarrow{R}_\ell,T) & \cdots & f_m(T,\overleftarrow{R}_\ell,T) \end{pmatrix}$$

Intuitively, the matrix D_T collects all the information of the degrees of the relations $R_1,\ldots,R_\ell$, while $\overleftarrow{D}_T$ the degrees of the reverse relations $\overleftarrow{R}_1,\ldots,\overleftarrow{R}_\ell$.

For two different types $S,T\in\mathcal{T}_\phi$, we define the matrix $D_{S\to T},\overleftarrow{D}_{S\to T}\in\mathbb{B}^{\ell\times m}$ as follows.

$$D_{S\to T} \ := \ \begin{pmatrix} f_1(S,R_1,T) & f_2(S,R_1,T) & \cdots & f_m(S,R_1,T) \\ f_1(S,R_2,T) & f_2(S,R_2,T) & \cdots & f_m(S,R_2,T) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(S,R_\ell,T) & f_2(S,R_\ell,T) & \cdots & f_m(S,R_\ell,T) \\ f_1(S,\overleftarrow{R}_1,T) & f_2(S,\overleftarrow{R}_1,T) & \cdots & f_m(S,\overleftarrow{R}_1,T) \\ f_1(S,\overleftarrow{R}_2,T) & f_2(S,\overleftarrow{R}_2,T) & \cdots & f_m(S,\overleftarrow{R}_2,T) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(S,\overleftarrow{R}_\ell,T) & f_2(S,\overleftarrow{R}_\ell,T) & \cdots & f_m(S,\overleftarrow{R}_\ell,T) \end{pmatrix}$$

and

$$\overleftarrow{D}_{S\to T} \ := \ \begin{pmatrix} f_1(T,\overleftarrow{R}_1,S) & f_2(T,\overleftarrow{R}_1,S) & \cdots & f_m(T,\overleftarrow{R}_1,S) \\ f_1(T,\overleftarrow{R}_2,S) & f_2(T,\overleftarrow{R}_2,S) & \cdots & f_m(T,\overleftarrow{R}_2,S) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(T,\overleftarrow{R}_\ell,S) & f_2(T,\overleftarrow{R}_\ell,S) & \cdots & f_m(T,\overleftarrow{R}_\ell,S) \\ f_1(T,R_1,S) & f_2(T,R_1,S) & \cdots & f_m(T,R_1,S) \\ f_1(T,R_2,S) & f_2(T,R_2,S) & \cdots & f_m(T,R_2,S) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(T,R_\ell,S) & f_2(T,R_\ell,S) & \cdots & f_m(T,R_\ell,S) \end{pmatrix}$$

Notice that in the matrix $D_{S\to T}$ the first ℓ rows contains the information on the degree of $R_1,\ldots,R_\ell$, and the last ℓ rows the information on the degree of $\overleftarrow{R}_1,\ldots,\overleftarrow{R}_\ell$ *from the type S to the type T*; while in the matrix $\overleftarrow{D}_{S\to T}$ it is the opposite and the direction is *from the type T to the type S*.

Now we define the formula $\mathsf{PREB}_\phi(x)$. First, we enumerate the set $\mathcal{T}_\phi=\{T_1,\ldots,T_n\}$ and the set $\mathcal{T}_\phi\times\mathcal{F}=\{(T_1,f_1),\ldots,(T_n,f_m)\}$. The formula $\mathsf{PREB}_\phi(x)$ is defined as follows.

$$\exists X_{(T_1,f_1)}\cdots\exists X_{(T_n,f_m)} \quad \left(x=\sum_{1\le i\le n}\ \sum_{1\le j\le m} X_{T_i,f_j}\right) \ \wedge\ \mathsf{PREB\text{-}Atom}_\phi(\bar{X}) \ \wedge\ \mathsf{CON}(\bar{X})$$

where $\bar{X}=(X_{(T_1,f_1)},\ldots,X_{(T_n,f_m)})$ the vector of all the variables $X_{(T,f)}$'s and

$$\mathsf{PREB\text{-}Atom}_\phi(\bar{X}) \ := \ \sum_{(T,f)\ \text{s.t.}\ \varphi\in T} X_{(T,f)}\ge k$$

and

$$\begin{aligned}\mathsf{CON}(\bar{X}) \ &:= \ \bigwedge_{T\in\mathcal{T}} \mathsf{Dir\text{-}REG\text{-}COMP}_{D_T,\overleftarrow{D}_T}(\bar{X}_T) \\ &\qquad \wedge \bigwedge_{1\le i\le n}\ \bigwedge_{1\le j\le i-1} \mathsf{BiREG\text{-}COMP}_{D_{T_i\to T_j},\overleftarrow{D}_{T_i\to T_j}}(\bar{X}_{T_i},\bar{X}_{T_j})\end{aligned}$$

where $\bar{X}_T = (X_{(T,f_1)}, \ldots, X_{(T,f_m)})$ and $\bar{X}_S = (X_{(S,f_1)}, \ldots, X_{(S,f_m)})$ the vector of variables associated with the types T and S, respectively.

Notice in the formula $\mathsf{BiREG\text{-}COMP}_{D_{T_i \to T_j}, \overleftarrow{D}_{T_i \to T_j}}(\bar{X}_{T_i}, \bar{X}_{T_j})$ the direction is always from T_i to T_j whenever $i \geq j+1$.

We claim that PREB_ϕ defines precisely the spectrum of ϕ. The proof is rather similar to the one in Proposition 6.1. We simply observe that the incoming R_i edges to an element v are precisely the outgoing $\overleftarrow{R}_i$ edges from v. Vice versa, the outgoing R_i edges from an element v are precisely the incoming $\overleftarrow{R}_i$ edges to v. The details can be found in Appendix H. ■

Similar to Proposition 6.2, one can show the following. Let $\phi \in \mathsf{QMLC}$ be a negation of a basic QMLC formula, say $\neg\exists^k \varphi$, where $\varphi \in \mathsf{MLC}$. Then there exists a Presburger formula $\mathsf{PREB}_\phi(x)$ such that $\textsc{Spec}(\phi) = \{n \mid \mathsf{PREB}_\phi(n) \text{ holds}\}$. We can achieve this by replacing the formula $\mathsf{PREB\text{-}Atom}$ above with $\sum_{(T,f) \text{ s.t. } \varphi \in T} X_{(T,f)} \leq k-1$.

To prove Theorem 2.1, we simply run through the same argument as in Theorem 6.3. Let $\phi \in \mathsf{QMLC}$. Push all the negations inside so that they are applied only to basic QMLC and redefine $\mathsf{PREB\text{-}Atom}_\phi$ inductively as follows.

- If $\phi := \exists^k \varphi$, then $\mathsf{PREB\text{-}Atom}_\phi := \sum_{(T,f) \text{ s.t. } \varphi \in T} X_{(T,f)} \geq k$.
- If $\phi := \neg\exists^k \varphi$, then $\mathsf{PREB\text{-}Atom}_\phi := \sum_{(T,f) \text{ s.t. } \varphi \in T} X_{(T,f)} \leq k-1$.
- If $\phi := \phi_1 \wedge \phi_2$, then $\mathsf{PREB\text{-}Atom}_\phi := \mathsf{PREB\text{-}Atom}_{\phi_1} \wedge \mathsf{PREB\text{-}Atom}_{\phi_2}$.
- If $\phi := \phi_1 \vee \phi_2$, then $\mathsf{PREB\text{-}Atom}_\phi := \mathsf{PREB\text{-}Atom}_{\phi_1} \vee \mathsf{PREB\text{-}Atom}_{\phi_2}$.

Now the formula $\mathsf{PREB}_\phi(x)$ is defined exactly like above.

$$\exists X_{(T_1,f_1)} \cdots \exists X_{(T_n,f_m)} \quad \left(x = \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq m} X_{T_i,f_j} \right) \wedge \ \mathsf{PREB\text{-}Atom}_\phi(\bar{X}) \ \wedge \ \mathsf{CON}(\bar{X})$$

where the formula CON is as in Proposition 8.1. The correctness of the formula $\mathsf{PREB}_\phi(x)$ can be verified by a straightforward induction. This completes our proof of Theorem 2.1.

9 Concluding remarks

By showing Presburger formulas for the existence of regular graphs, we have shown that the spectra of $\mathsf{FO}^2\mathsf{C}$ formulae are constructive semilinear sets. From our proof, it can be immediately deduced that the many-sorted spectra of $\mathsf{FO}^2\mathsf{C}$ are also semilinear. As far as our knowledge is concerned, the logic $\mathsf{FO}^2\mathsf{C}$ is the first logic whose spectra is closed under complement without any restriction on the vocabulary nor in the interpretation. The semilinearity of its spectra also gives us an interesting insight on the nature of two-variable logic – that each of its models is simply a collection of regular graphs.

Another interesting open question is, how can be $\mathsf{FO}^2\mathsf{C}$ extended while keeping decidability? Using three variables ($\mathsf{FO}^3\mathsf{C}$) one can easily encode a grid; therefore, the Parikh images no longer are decidable. However, we could extend $\mathsf{FO}^2\mathsf{C}$ by giving access to a relation having a property which is undefinable in $\mathsf{FO}^2\mathsf{C}$, such as transitivity. In particular, $\mathsf{FO}^2\mathsf{C}(<)$, that is, the logic $\mathsf{FO}^2\mathsf{C}$ with access to a total order on the universe, seems powerful: Petri net reachability [29, 24, 25] reduces to Parikh image membership for $\mathsf{FO}^2\mathsf{C}(<)$. We do not know whether a reduction exists in the other direction. Another possible extension is to add an equivalence relation to $\mathsf{FO}^2\mathsf{C}$.

References

[1] G. Asser. *Das Repräsentenproblem in Prädikatenkalkül der ersten Stufe mit Identität.* Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 1: 252–263 (1955).

[2] A. Durand, N. D. Jones, J. A. Makowsky, M. More. *Fifty years of the spectrum problem: survey and new results.* Bulletin of Symbolic Logic 18(4): 505–553 (2012).

[3] H.-D. Ebbinghaus, J. Flum. *Finite model theory.* Perspectives in Mathematical Logic, Springer 1995.

[4] R. Fagin. *Contributions to the model theory of finite structures.* Ph.D. thesis, University of California, Berkeley, 1973.

[5] R. Fagin. *Finite model theory-a personal perspective.* Theoretical Computer Science 116: 3–31 (1993).

[6] W. Fan, L. Libkin. *On XML integrity constraints in the presence of DTDs.* Journal of the ACM 49(3): 368–406 (2002).

[7] E. Fischer, J. A. Makowsky. *On spectra of sentences of monadic second order logic with counting.* Journal of Symbolic Logic 69(3):617–640 (2004).

[8] S. Ginsburg, E. H. Spanier. *Semigroups, Presburger Formulas, and Languages.* Pacific Journal of Mathematics 16:285–296 (1966).

[9] É. Grandjean. *The spectra of first-order sentences and computational complexity.* SIAM Journal on Computing 13(2): 356–373 (1984).

[10] É. Grandjean. *Universal quantifiers and time complexity of random access machines.* Mathematical Systems Theory 18(2): 171–187 (1985).

[11] É. Grandjean. *First-order spectra with one variable.* Journal of Computer and System Sciences 40(2): 136–153 (1990).

[12] E. Grädel. *Why are Modal Logics so Robustly Decidable?* Current Trends in Theoretical Computer Science 2001.

[13] E. Grädel, M. Otto. *On Logics with Two Variables.* Theoretical Computer Science 224(1-2): 73–113 (1999).

[14] E. Grädel, M. Otto, E. Rosen. *Two-Variable Logic with Counting is Decidable.* In LICS 1997.

[15] M. Grohe, J. A. Makowsky, Editors. *Model Theoretic Methods in Finite Combinatorics.* Contemporary Mathematics 558, American Mathematical Society, 2011.

[16] Y. Gurevich and S. Shelah. *Spectra of monadic second-order formulas with one unary function.* In LICS 2003.

[17] S. Hakimi. *On realizability of a set of integers as degrees of the vertices of a linear graph. I.* Journal of the Soc. Indust. Appl. Math. 10(3): 496–506 (1962).

[18] A. Hunter. *Spectrum hierarchies and subdiagonal functions.* In LICS 2003.

[19] N. Immerman. DSPACE[n^k] = VAR[k+1]. In *Structure in Complexity Theory* Conference 1991.

[20] N. Immerman. *Descriptive Complexity: A Logician's Approach to Computation.* Notices of the AMS 42:10 (1995).

[21] N. Immerman. *Descriptive complexity.* Graduate texts in computer science, Springer 1999.

[22] N. Jones, A. Selman. *Turing machines and the spectra of first-order formulas.* The Journal of Symbolic Logic, 39: 139–150 (1974). (Also in STOC 1972.)

[23] H. Landau. *On dominance relations and the structure of animal societies. III. The condition for a score structure.* Bulletin of Mathematical Biophysics 15(2): 143–148 (1953).

[24] S. Kosaraju. Decability of reachability in vector addition systems. In STOC 1982.

[25] J. Leroux. The General Vector Addition System Reachability Problem by Presburger Inductive Invariants. In LICS 2009.

[26] L. Libkin. *Elements of Finite Model Theory.* Springer 2004.

[27] J. Lynch. *Complexity classes and theories of finite models. Mathematical Systems Theory* 15(2): 127–144 (1982).

[28] C. Lutz, U. Sattler, F. Wolter. *Modal logic and the two-variable fragment.* In CSL 2001.

[29] E. Mayr. An algorithm for the general petri net reachability problem. In STOC 1981.

[30] M. Mortimer. *On languages with two variables.* Mathematical Logic Quarterly, 21(1): 135–140 (1975).

[31] M. Otto. *Bounded Variable Logics and Counting - A Study in Finite Models* Lecture Notes in Logic volume 9, Springer 1997.

[32] L. Pacholski, W. Szwast, L. Tendera. *Complexity Results for First-Order Two-Variable Logic with Counting.* SIAM Journal on Computing 29(4): 1083–1117 (2000).

[33] D. Rödding, H. Schwichtenberg. *Bemerkungen zum Spektralproblem.* Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 18: 1–12 (1972).

[34] H. Scholz. *Ein ungelöstes Problem in der symbolischen Logic.* The journal of Symbolic Logic, 17: 160 (1952).

[35] K. N. Verma, H. Seidl, T. Schwentick. *On the Complexity of Equational Horn Clauses.* In CADE 2005.

[36] M. Vardi. *Why is Modal Logic So Robustly Decidable?* In Descriptive Complexity and Finite Models 1996.

[37] A. Woods. *Some Problems in Logic and Number Theory and their Connections.* Ph.D. thesis, University of Manchester, 1981.

[38] C. Wrathall. *Rudimentary Predicates and Relative Computation.* In SIAM Journal on Computing 7(2): 194–209 (1978).

APPENDICES

A The equivalence between $\mathsf{FO}^2\mathsf{C}$ and QMLC

Let $\phi \in \mathsf{FO}^2\mathsf{C}$ and let $\mathcal{R}$ and $\mathcal{S}$ be the set of binary and unary predicates appearing in ϕ, respectively. The construction of the QMLC formula for ϕ is as follows.

1. Convert the sentence ϕ into its "normal form" ϕ'.[‖]
2. Convert the sentence ϕ' into a "quantified modal logic" (QMLC) sentence φ such that for all structure $\mathfrak{A}$, we have $\mathfrak{A} \models \phi'$ if and only if $\mathfrak{A} \models \varphi$.

We remark these steps are simply adaptation of the technique in [28].

The sentence $\phi \in \mathsf{FO}^2\mathsf{C}$ is in *normal form*, if all the quantifiers are either of form

$$\exists^k y \ \Big(R(x,y) \wedge \psi(y)\Big),$$

or of form

$$\exists^k x \ \psi(x)$$

and all other applications of variables are of form $a_i(x)$.

We claim that every sentence $\phi \in \mathsf{FO}^2\mathsf{C}$ can be converted into its equivalent sentence in normal form. It can be done as follows.

- First, we rewrite every subformula of the form $\exists^k y \ \psi(x,y)$ with one free variable x into the following form:

$$\begin{array}{c} \psi(x,x) \ \wedge \ \exists^{k-1} y \ \Big((x \neq y) \wedge \psi(x,y)\Big) \\ \vee \\ \exists^k y \ \Big((x \neq y) \wedge \psi(x,y)\Big) \end{array}$$

  After such rewriting, we can assume that every quantifier in ϕ is of the form $\exists^k y \ ((x \neq y) \wedge \psi(x,y))$.

- Second, every quantification $\exists^k y \ ((x \neq y) \wedge \psi(x,y))$, in which $\psi(x,y)$ contains a subformula $\alpha(x)$ depending only on x, can be rewritten into the form:

$$\begin{array}{c} \neg\alpha(x) \ \wedge \ \exists^k y \ \Big((x \neq y) \wedge \psi_0(x,y)\Big) \\ \vee \\ \alpha(x) \ \wedge \ \exists^k y \ \Big((x \neq y) \wedge \psi_1(x,y)\Big) \end{array}$$

  where $\psi_0(x,y)$ and $\psi_1(x,y)$ are obtained from ψ by replacing $\alpha(x)$ with false and true, respectively. We can repeat this until $\psi(x,y)$ no longer has a subformula depending only on x.
  After such rewriting we can assume that every quantifier in ϕ is of the form

$$\exists^k y \ ((x \neq y) \wedge \psi(x,y)),$$

  where $\psi(x,y)$ does not contain any subformula depending only on x.

[‖] We would like to remark that the normal form here is different from the standard Scott's normal form.

- Third, every quantification $\exists^k y \Big((x \neq y) \wedge \psi(x,y)\Big)$ can be rewritten into the form:

$$\bigvee_{f \in \Delta^k_{\mathcal{R}}} \bigwedge_{R \in \mathcal{R}} \exists^{f(R)} y \Big(R(x,y) \wedge \psi_R(y)\Big)$$

  where $\Delta^k_{\mathcal{R}}$ is the set of all function $f : \mathcal{R} \to \mathbb{N}$ such that $\sum_{R \in \mathcal{R}} f(R) = k$, and $\psi_R(y)$ is obtained from $\psi(x,y)$ by replacing each $R'(x,y)$ with true if $R = R'$, and false otherwise.

By performing these three steps, we get an $\mathsf{FO}^2\mathsf{C}$ formula in the normal form.

Now given an $\mathsf{FO}^2\mathsf{C}$ sentence ϕ in its normal form, the construction of its QMLC sentence $\psi = F(\phi)$ can be done inductively as follows. There are two cases.

1. ϕ has no free variable.
   - If ϕ is $\neg\phi_1$, then $F(\phi) = \neg F(\phi_1)$.
   - If ϕ is $\phi_1 \wedge \phi_2$, then $F(\phi) = F(\phi_1) \wedge F(\phi_2)$.
   - If ϕ is $\exists^k x\ \phi_1(x)$, then $F(\phi) = \exists^k F(\phi_1(x))$.
2. ϕ has one free variable x.
   - If $\phi(x)$ is $S(x)$, then $F(\phi(x)) = S$.
   - If $\phi(x)$ is $\phi_1(x) \wedge \phi_2(x)$, then $F(\phi(x)) = F(\phi_1(x)) \wedge F(\phi_2(x))$.
   - If $\phi(x)$ is $\neg\phi_1(x)$, then $F(\phi(x)) = \neg F(\phi_1(x))$.
   - If $\phi(x)$ is $\exists^k y\ R(x,y) \wedge \psi(x,y)$, then $F(\phi(x)) = \Diamond^k_R F(\psi(x,y))$.

   The case when ϕ has one free variable y can be handled in a symmetrical way.

By a straightforward induction, we can show that for every structure $\mathfrak{A}$, $\mathfrak{A} \models \phi$ if and only if $\mathfrak{A} \models F(\phi)$. This concludes the conversion from the normal forms of $\mathsf{FO}^2\mathsf{C}$ to QMLC formulae.

B Proof of Theorem 4.1

For convenience, we repeat Theorem 4.1 here.

Theorem 4.1. *For every $D \in \mathbb{B}^{\ell \times m}$, there exists a Presburger formula $\mathsf{REG}_D(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. There exists a D-regular ℓ-type graph of size $\bar{N}$ if and only if $\mathsf{REG}_D(\bar{N})$ holds.*

The essence of the theorem lies on the case when $D \in \mathbb{N}^{\ell \times m}$. This is what we will consider first for the sake of presentation. The proof is divided into three successive steps.

(1) First, we consider the simplest case when $D \in \mathbb{N}^{1 \times 1}$ in Subsection B.1.

(2) Then we generalise it to the case $D \in \mathbb{N}^{1 \times m}$ in Subsection B.2. This is done by induction on m.

(3) Finally we consider the case of arbitrary $D \in \mathbb{N}^{\ell \times m}$ in Subsection B.3. This is done by induction on ℓ.

When D contains an element of $\mathbb{B}$, it will be easier since there is more freedom in adding more edges. However, the proof will be more tedious. For this reason, we postpone the case when D contains an element of $\mathbb{B}$ until Subsection B.4.

B.1 When $D \in \mathbb{N}^{1\times 1}$

In this subsection we deal with the case when D consists of only one entry $d \in \mathbb{N}$. This means that the degree of every vertex is the same. For convenience, we will simply write d-regular graphs, instead of a more cumbersome notation (d)-regular graph.

We have the following proposition which appears to be a folklore in the graph theory. We present the proof here since we are going to need it.

Proposition B.1 *Let $d \geq 0$. Then, for every $N \geq 0$, the following holds.*

(a) There exists a d-regular graph of size N if and only if $N \geq d+1$ and the product dN is an even number.

(b) If the product Nd is an odd number and $N \geq d+1$, then there is a graph of size N in which there is exactly one vertex with degree $d-1$, and all the other $N-1$ vertices have degree d.

Proof. Let $d \geq 0$. We first prove item (a). The "only if" direction is straightforward. If a d-regular graph has size N, then $N \geq d+1$. Moreover, the product dN is precisely twice the number of edges, hence, an even number.

The "if" direction is as follows. Let $N \geq d+1$ and the product dN is an even number. There are two cases.

- If d is even, then we construct a d-regular graph $G = (V, E)$ with N vertices as follows. Let $V = \{v_0, \ldots, v_{N-1}\}$. For each $v_i \in V$, we put an edge between v_i and each of the vertices $v_{i-\frac{d}{2}}, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{i+\frac{d}{2}}$. (Here the minus and plus are on modulo N.) It is obvious then that every vertex has degree d.
- If d is odd, then N must be even. The construction of a d-regular graph of size N is as follows. First, we construct a $(d-1)$-regular graph $G' = (V, E')$ of size N. (Since d is odd, then $(d-1)$ must be even.) Now every vertex has degree $(d-1)$. To increase the degree of each vertex by one, we can do the *pairing* between two vertices as follows. Let $V = \{v_0, \ldots, v_{N-1}\}$. For each $0 \leq i \leq N/2 - 1$, we add a new edge between the vertex v_i and its *partner* $v_{N/2+i}$. (Note that N is even in this case, hence every vertex has its partner.) This way we obtain a d-regular graph of size N.

Next, we prove item (b). If Nd is odd, then both N and d are odd. Now we perform the construction of d-regular graph of size N as in the case of odd d in item (a). But when we do the pairing, there is one vertex without a partner. This vertex has degree $d-1$, while the rest have degree d each. ■

B.2 When $D \in \mathbb{N}^{1\times m}$

Next, we consider the case when $\ell = 1$ and $m \geq 1$, that is, when D consists of only one row vector $\bar{d}$, or in other words, the graph is a 1-type graph. The proof is by induction on m with the basis $m = 1$ already settled in Proposition B.1. Since we only consider the case of 1-type graph, we will simply write $\deg(u)$ to denote the degree of the vertex u, instead of $\deg_{E_1}(u)$.

We need the following notion of *almost $\bar{d}$-regular* graph. For a vector $\bar{d} \in \mathbb{N}^m$, a 1-type graph $G = (V, E)$ is *almost $\bar{d}$-regular* graph, if there is a partition $V_1 \cup \cdots \cup V_m$ of its vertices such that there exists exactly one $j \in \{1, \ldots, m\}$, where

- for all $i \neq j$, for every vertex $u \in V_i$, $\deg(u) = d_i$;
- there is exactly one vertex $u \in V_j$ where $\deg(u) = d_j - 1$, and all the other vertices in V_j have degree d_j.

The special vertex in V_j with degree $d_j - 1$ is called the *culprit* vertex. We say that the graph is of size $\bar{N} = (N_1, \ldots, N_m)$, where $(N_1, \ldots, N_m) = (|V_1|, \ldots, |V_m|)$. Similar to the above, we call the partition $V_1 \cup \cdots \cup V_m$ *a partition of the almost $\bar{d}$-regularity* of the graph G.

In the following $\bar{1}$ denotes the vector $(1, \ldots, 1) \in \mathbb{N}^m$, that is, $\bar{1}$ is a vector whose components are all one. Lemma B.2 below is a generalisation of Proposition B.1.

Lemma B.2 *Let $\bar{d} \in \mathbb{N}^m$ such that it does not contain zero entry. Then, for every $\bar{N}$ such that $\bar{N} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + (\bar{d} \cdot \bar{1})$, the following holds.*

(a) If the dot product $\bar{d} \cdot \bar{N}$ is even, then there is a $\bar{d}$-regular graph of size $\bar{N}$.

(b) If the dot product $\bar{d} \cdot \bar{N}$ is odd, then there is an almost $\bar{d}$-regular graph of size $\bar{N}$.

Proof. The proof is by induction on m. The base case $m = 1$ has been established in parts (a) and (b) of Propositions B.1, since $d^4 + 1 \geq d + 1$. For the induction hypothesis, we assume that the lemma holds for the case of m. For the induction step, we are going to establish the case of $m + 1$.

Let $\bar{d} = (d_1, \ldots, d_{m+1}) \in \mathbb{N}^{m+1}$. Now suppose $\bar{N} = (N_1, \ldots, N_{m+1})$ such that $\bar{N} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + \bar{d} \cdot \bar{1}$.

We first prove part (a). Suppose $\bar{d} \cdot \bar{N}$ is even. There are two cases.

Case 1: For all $i \in \{1, \ldots, m + 1\}$, $N_i \geq d_i + 1$.
In this case, by Proposition B.1, we construct the following graph G_i, for each $i \in \{1, \ldots, m + 1\}$.

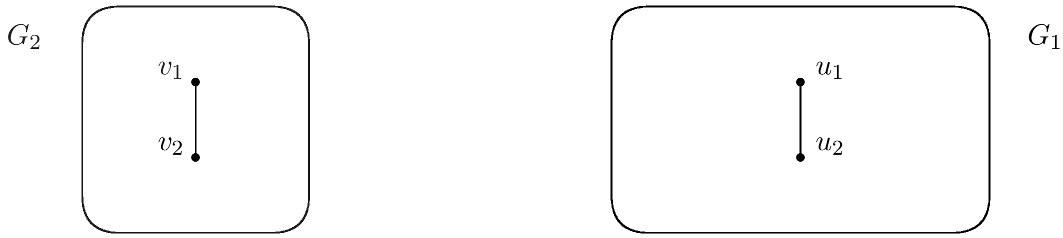
- If $N_i d_i$ is an even number, the graph G_i is a d_i-regular graph of size N_i.
- If $N_i d_i$ is an odd number, the graph G_i is an almost d_i-regular graph of size N_i.

Since $\bar{N} \cdot \bar{d}$ is an even number, there are even number of almost d_i-regular graphs from all the $G_1, \ldots, G_{m+1}$. Hence, there are even number of culprit vertices, say $2k$. We can form k number of pairs of culprit vertices and connect each pair with an edge. This way we obtain a $\bar{d}$-regular graph of size $\bar{N}$.

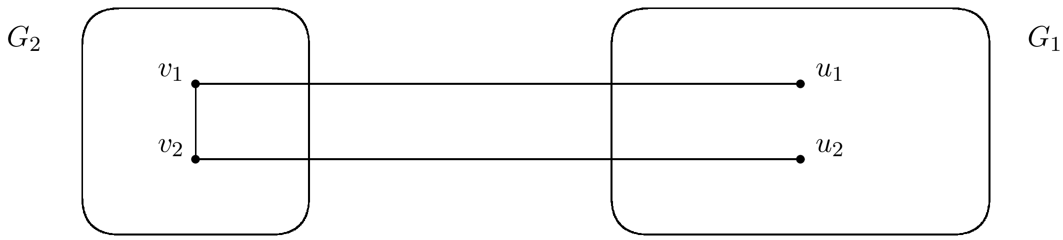
Case 2: There exists $i \in \{1, \ldots, m + 1\}$ such that $1 \leq N_i \leq d_i$.
Without loss of generality, we assume that $N_{m+1} \leq d_{m+1}$. There are two cases:

- Both $d_{m+1} N_{m+1}$ and $\sum_{i=1}^{m} d_i N_i$ are even.
Since $\bar{N} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + \bar{d} \cdot \bar{1}$, we have $\sum_{i=1}^{m} N_i \geq (\sum_{i=1}^{m} d_i)^4 + \sum_{i=1}^{m} d_i$, hence we can apply the induction hypothesis, from which we obtain a $(d_1, \ldots, d_m)$-regular graph G_1 of size $(N_1, \ldots, N_m)$. Let G_2 be a clique of N_{m+1} nodes. We are going to combine G_1 and G_2 to obtain a $\bar{d}$-regular graph of size $\bar{N}$. The idea is to increase the degree of each node in G_2, while preserving the degree of each node in G_1.
Each vertex in G_2 has degree $N_{m+1} - 1$. We have to increase it to d_{m+1}. So we need $(d_{m+1} - N_{m+1} + 1)$ additional edges adjacent to each vertex in G_2, or $N_{m+1}(d_{m+1} - N_{m+1} + 1)$ edges adjacent to G_2 in total.
Intuitively, we do the following. For each edge $(u_1, u_2) \in E(G_1)$, we choose two vertices $v_1, v_2 \in V(G_2)$, delete the edge (u_1, u_2) and add in the edges $(v_1, u_1), (v_2, u_2)$.

We call the edges $(v_1, u_1), (v_2, u_2)$ *crossing edges* between G_1 and G_2. See, for example, the graph below. The graph G_2 is a clique of d_{m+1} vertices, while G_1 is the $(d_1, \ldots, d_m)$-regular graph of size $(N_1, \ldots, N_m)$. Now we pick an edge (u_1, u_2) in G_1 and two vertices v_1, v_2 in G_2 as illustrated below.



From the above graph, we "omit" the edge (u_1, u_2) and "add in" the two crossing edges (v_1, u_1) and (v_2, u_2). See the illustration below. This way we increase the degrees of both v_1, v_2 by 1, while the degrees of both u_1, u_2 stay the same.



We do this repeatedly until the degree of each vertex in G_2 becomes d_{m+1}. There are two things to prove here: (1) that it is possible to arrive in the situation that the degree of each vertex in G_2 becomes d_{m+1}; (2) there are sufficient edges in G_2 for that.

First, we prove (1). The initial total degree of vertices in G_2 is $N_{m+1}(N_{m+1} - 1)$, which is an even number. Each time we perform the operation above, we increase the total degree of vertices in G_2 by two. Since $N_{m+1}d_{m+1}$ is even, and hence so is the number $N_{m+1}(d_{m+1} - N_{m+1} + 1)$, at the end we will reach the situation that each vertex in G_2 becomes d_{m+1}.

Now we prove (2). We can pick $\frac{N_{m+1}(d_{m+1}-N_{m+1}+1)}{2}$ disjoint edges in G_1. (Note that $N_{m+1}(d_{m+1} - N_{m+1} + 1)$ is an even number.) Those edges exist for the following reason. Since $\bar{N} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + \bar{d} \cdot \bar{1}$ and $N_{m+1} \leq d_{m+1}$, we have $\sum_{1 \leq i \leq m} N_i \geq (\bar{d} \cdot \bar{1})^4$. Each of the vertex in G_1 has degree $< \bar{d} \cdot \bar{1}$. Hence, there are

$$\frac{\sum_{1 \leq i \leq m} N_i}{(\bar{d} \cdot \bar{1})^2} \quad \geq \quad d_{m+1}^2 \quad \geq \quad \frac{N_{m+1}(d_{m+1} - N_{m+1} + 1)}{2}$$

vertices with disjoint neighbours.

- Both $d_{m+1}N_{m+1}$ and $\sum_{i=1}^{m} d_i N_i$ are odd.
  Applying the induction hypothesis, we obtain an almost $(d_1, \ldots, d_m)$-regular graph

G_1 of size $(N_1, \ldots, N_m)$. Let u be the culprit vertex in G_1. Let G_2 be a clique of N_{m+1} nodes. Note that the sum of the degrees of vertices in G_2 is $N_{m+1}(N_{m+1}-1)$. We perform similar operation as in the case above. For two vertices $v_1, v_2 \in V(G_2)$ whose degrees are $\leq d_{m+1} - 1$, or a vertex $v_1 = v_2$ whose degree is $\leq t_{m+1} - 2$, we pick an edge $(u_1, u_2) \in E(G_2)$, delete it and add two crossing edges (v_1, u_1) and (v_2, u_2).

We stop the process until there is only one vertex $v \in V(G_2)$ whose degree is $d_{m+1} - 1$ and all other vertices in $V(G_2)$ have degree d_{m+1}. Since $N_{m+1}d_{m+1}$ is odd number, while $N_{m+1}(N_{m+1} - 1)$ is even number, such vertex v should exist at the end of the process of "converting edges in G_1 to crossing edges." That there are enough edges in $E(G_1)$ is already established above. Now, connecting the culprit vertices u and v with an edge, we obtain $\bar{d}$-regular graph of size $\bar{N}$.

Now we prove part (b). The proof is very similar to part (a). The only difference is that there should be a culprit vertex.

There are two cases.

Case 1: For all $i \in \{1, \ldots, m+1\}$, $N_i \geq d_i + 1$.

In this case, by Proposition B.1, we construct the following graph G_i, for each $i \in \{1, \ldots, m+1\}$.

- If $N_i d_i$ is an even number, the graph G_i is a d_i-regular graph of size N_i.
- If $N_i d_i$ is an odd number, the graph G_i is an almost d_i-regular graph of size N_i.

Since $\bar{N} \cdot \bar{d}$ is an odd number, there are odd number of almost d_i-regular graphs from all the $G_1, \ldots, G_{m+1}$. Hence, there are odd number of culprit vertices, say $2k+1$. From the $2k$ culprit vertices, we can form k number of pairs of culprit vertices and connect each pair with an edge. Thence, we are left with one culprit vertex. This way we obtain an almost $\bar{d}$-regular graph of size $\bar{N}$.

Case 2: There exists $i \in \{1, \ldots, m+1\}$ such that $N_i \leq d_i$.

Without loss of generality, we assume that $N_{m+1} \leq d_{m+1}$. There are two cases.

- $d_{m+1}N_{m+1}$ is odd and $\sum_{i=1}^{m} d_i N_i$ is even.
  Applying the induction hypothesis, we obtain a $(d_1, \ldots, d_m)$-regular graph G_1 of size $(N_1, \ldots, N_m)$. Let G_2 be a clique of size N_{m+1}. By converting an edge in $E(G_1)$ into two crossing edges to increase the degrees of vertices in G_2, we can get the degrees of all vertices in G_2, but one, are d_{m+1}, and one vertex has degree $d_{m+1} - 1$.
- $d_{m+1}N_{m+1}$ is even and $\sum_{i=1}^{m} d_i N_i$ is odd.
  Applying the induction hypothesis, we obtain an almost $(d_1, \ldots, d_m)$-regular graph G_1 of size $(N_1, \ldots, N_m)$. Let G_2 be a clique of size N_{m+1}. By converting an edge in $E(G_1)$ into two crossing edges to increase the degrees of vertices in G_2, we can get the degrees of all vertices in G_2 to be d_{m+1}.

This competes the proof of Lemma B.2. ■

Lemma B.2 can then be easily extended to obtain a Presburger formula for the existence of $\bar{d}$-regular graphs when $\bar{d} \in \mathbb{N}^m$, as stated in Theorem B.3 below.

Theorem B.3 *For every $\bar{d} \in \mathbb{N}^m$, there exists a Presburger formula $\mathsf{REG}_{\bar{d}}(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. There exists a $\bar{d}$-regular graph of size $\bar{N}$ if and only if $\mathsf{REG}_{\bar{d}}(\bar{N})$ holds.*

Proof. The proof is a direct application of Lemma B.2. We assume that all the entries in $\bar{d}$ are not zero. We define the following set I.

$$I \ := \ \{\bar{N} \mid \bar{N} \cdot \bar{1} < (\bar{d} \cdot \bar{1})^4 + (\bar{d} \cdot \bar{1}) \text{ and there exists a } \bar{d}\text{-regular graph of size } \bar{N}\}$$

Such set can be computed since the number of $\bar{N}$ such that $\bar{N} \cdot \bar{1} \leq (\bar{d} \cdot \bar{1})^4 + (\bar{d} \cdot \bar{1})$ is finite and for each $\bar{N}$ there are only finitely many graphs whose size is $\bar{N}$. Now, we define the formula $\mathsf{REG}_{\bar{d}}(\bar{X})$ to be the following formula.

$$\Big(\bar{X} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + (\bar{d} \cdot \bar{1}) \wedge (\exists z \ \bar{X} \cdot \bar{d} = 2z)\Big) \ \vee \ \bigvee_{\bar{N} \in I} \bar{X} = \bar{N}$$

The formula is a Presburger formula since $\bar{d}$ are constants. That $\mathsf{REG}_{\bar{d}}(\bar{X})$ is the desired formula follows from Lemma B.2 and the fact that if G is a $\bar{d}$-regular graph of size $\bar{N}$, then $\bar{N} \cdot \bar{d}$ is precisely twice the number of edges, hence an even number.

For the case when $\bar{d}$ contains zero entries, we do the following. Suppose $\bar{d} = (d_1, \ldots, d_m)$ and let $J = \{i \mid d_i = 0\}$

$$\mathsf{REG}_{\bar{d}}(\bar{X}) \ := \ \bigwedge_{i \in J} \ X_i \geq 0 \ \wedge \ \mathsf{REG}_{\bar{d}'}(\bar{X}'),$$

where $\bar{d}'$ and $\bar{X}'$ are the vectors $\bar{d}$ and $\bar{X}$ without the entries in J, respectively, and $\mathsf{REG}_{\bar{d}'}(\bar{X}')$ is defined as above. ■

B.3 When $D \in \mathbb{N}^{\ell \times m}$

In this section we consider the case when $D \in \mathbb{N}^{\ell \times m}$. Let $D = \begin{pmatrix} \bar{d}_1 \\ \vdots \\ \bar{d}_\ell \end{pmatrix} \in \mathbb{N}^{\ell \times m}$, where $\bar{d}_1, \ldots, \bar{d}_\ell$ are the row vectors of D.

We define the following set

$$I_D \ := \ \{\bar{N} \mid \bar{N} \cdot \bar{1} < 2\ell(D \cdot \bar{1})^2 + 3\ell \text{ and there exists a } C\text{-regular graph of size } \bar{N}\}$$

Then, we define the formula $\mathsf{REG}_D(\bar{X})$ inductively, where $\bar{X} = (X_1, \ldots, X_m)$ as follows. For simplicity, we assume that D does not contain zero column. (If column i in D is a zero column, we simply add $X_i \geq 0$ and ignore the column.)

- When $\ell = 1$:

$$\mathsf{REG}_D(\bar{X}) \ := \ \mathsf{REG}_{\bar{d}_1}(\bar{X})$$

- When $\ell \geq 2$:

$$\mathsf{REG}_D(\bar{X}) \ := \ \bigvee_{\bar{N} \in I_D} \bar{X} = \bar{N} \quad \vee \quad \bigvee_{1 \leq j \leq \ell} \Big(\bar{X} \cdot \chi(\bar{d}_j) \geq 2(D \cdot \bar{1})^2 + 3 \quad \wedge \quad \mathsf{REG}_{D - \bar{d}_j}(\bar{X}) \quad \wedge \quad \mathsf{REG}_{\bar{d}_j}(\bar{X}) \Big)$$

where $\chi(\bar{d}_j)$ is the characteristic vector of $\bar{d}_j$** and $D-\bar{d}_j$ denotes the matrix D without row j.

We are going to prove that for every $D \in \mathbb{N}^{\ell\times m}$, there exists a D-regular graph of size $\bar{N}$ if and only if the statement $\mathsf{REG}_D(\bar{N})$ holds. The proof is by induction on ℓ. The basis $\ell = 1$ has been established in Theorem B.3. For the induction step, we assume that it holds for the case of $\ell-1$ and we are going to prove the case ℓ.

We first prove the "only if" direction. Suppose $G = (V, E_1, \ldots, E_\ell)$ is D-regular of size $\bar{N}$. If $\bar{N} \in I_D$, then $\mathsf{REG}_D(\bar{N})$ holds. So suppose $\bar{N} \notin I_D$, that is, $\bar{N}\cdot\bar{1} \geq 2\ell(D\cdot\bar{1})^2+3\ell$. Since D does not contain zero column, there exists $j \in \{1,\ldots,\ell\}$ such that $\bar{N}\cdot\chi(\bar{d}_j) \geq 2(D\cdot\bar{1})^2+3$. Moreover, if $G = (V, E_1, \ldots, E_\ell)$ is D-regular of size $\bar{N}$, then G is also $(D-\bar{d}_j)$-regular and $\bar{d}_j$-regular. By the induction hypothesis, both $\mathsf{REG}_{D-\bar{d}_j}(\bar{N})$ and $\mathsf{REG}_{\bar{d}_j}(\bar{N})$ hold.

We now prove the "if" direction. Suppose $\mathsf{REG}_D(\bar{N})$ holds. If $\bar{N} \in I_D$, then there exists a D-regular graph of size $\bar{N}$ and we are done. So suppose $\bar{N} \notin I_D$. Hence there exists $j \in \{1,\ldots,\ell\}$ such that

$$\bar{N}\cdot\chi(\bar{d}_j) \geq 2(D\cdot\bar{1})^2+3 \quad\wedge\quad \mathsf{REG}_{D-\bar{d}_j}(\bar{N}) \quad\wedge\quad \mathsf{REG}_{\bar{d}_j}(\bar{N})$$

For simplicity, we assume that $j = \ell$. By induction hypothesis, there exists a $(D-\bar{d}_\ell)$-regular graph $G_1 = (V_1, E_1, \ldots, E_{\ell-1})$ of size $\bar{N}$, and by definition, $E_1, \ldots, E_{\ell-1}$ are pairwise disjoint. By Theorem B.3, there exists a $\bar{d}_\ell$-regular graph $G_2 = (V_2, E_\ell)$ of size $\bar{N}$. We can assume that $V_1 = V_2$ since G_1 and G_2 are of the same size $\bar{N}$.

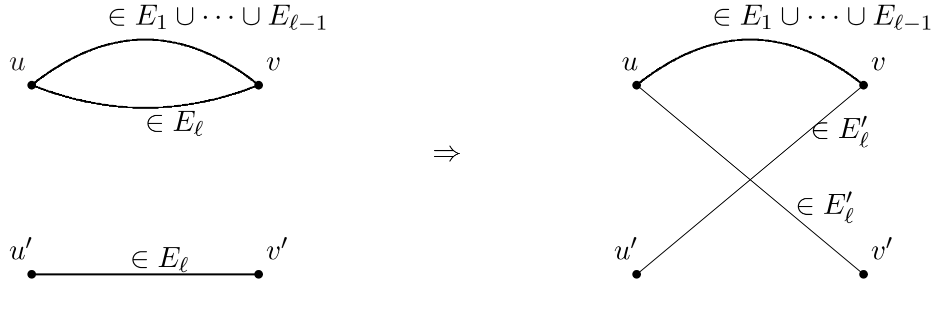
We are going to combine G_1 and G_2 into one graph to get an ℓ-type D-regular graph $G_1 = (V_1, E_1, \ldots, E_\ell)$ of size $\bar{N}$. If $E_\ell\cap(E_1\cup\cdots\cup E_{\ell-1}) = \emptyset$, then the graph $G = (V, E_1, \ldots, E_\ell)$ is the desired D-regular (ℓ)-type graph of size $\bar{N}$, and we are done.

Now suppose $E_\ell \cap (E_1 \cup \cdots \cup E_{\ell-1}) \neq \emptyset$. We are going to construct another graph $G_2' = (V, E_\ell')$ such that

$$|E_\ell' \cap (E_1 \cup \cdots \cup E_{\ell-1})| \quad<\quad |E_\ell \cap (E_1 \cup \cdots \cup E_{\ell-1})|$$

We do this repeatedly until at the end we obtain a graph $G_2'' = (V, E_\ell'')$ such that $E_\ell'' \cap (E_1 \cup \cdots \cup E_{\ell-1}) = \emptyset$.

Let $(u, v) \in E_\ell \cap (E_1 \cup \cdots \cup E_{\ell-1})$. The number of vertices reachable in from u and v within distance 2 (by any of edges in $E_1, \ldots, E_\ell$) is bounded from above by $2(D\cdot\bar{1})^2+2$. Since $\bar{N}\cdot\chi(\bar{d}_\ell) \geq 2(D\cdot\bar{1})^2+3$, there exists $(u', v') \in E_\ell$ such that $(u, u'), (v, v') \notin E_1 \cup\cdots\cup E_\ell$. See the left side of the illustration below.



**The characteristic vector of $\bar{c} = (c_1, \ldots, c_m) \in \mathbb{N}^m$ is defined as $\chi(c_1, \ldots, c_m) := (b_1, \ldots, b_m) \in \{0,1\}^m$ where $b_i = 0$ if $c_i = 0$, and $b_i = 1$ if $c_i \geq 1$.

Now we define E'_ℓ by deleting the edges $(u,v),(u',v')$ from E_ℓ, while adding the edges $(u,v'),(u',v)$ into E_ℓ. Formally,

$$E'_\ell \;:=\; (E_\ell - \{(u,v),(u',v')\}) \cup \{(u,u'),(v,v')\}$$

See the right side of the illustration above.

Now it is straightforward that $G' = (V, E'_\ell)$ is still a $\bar{d}_\ell$-regular graph of size $\bar{N}$, while

$$|E'_\ell \cap (E_1 \cup \cdots \cup E_{\ell-1})| \;<\; |E_\ell \cap (E_1 \cup \cdots \cup E_{\ell-1})|$$

We perform this operation until $E_{\ell+1} \cap (E_1 \cup \cdots \cup E_\ell) = \emptyset$. This completes our proof when $D \in \mathbb{N}^{\ell \times m}$.

B.4 When D contains elements from ${}^{\blacktriangleright}\mathbb{N}$

In this subsection we consider the case when D contains elements from ${}^{\blacktriangleright}\mathbb{N}$. Recall that $\mathbb{B} = \mathbb{N} \cup {}^{\blacktriangleright}\mathbb{N}$.

We are going to use the following notations. For ${}^{\blacktriangleright}d \in {}^{\blacktriangleright}\mathbb{N}$, we write $\lfloor {}^{\blacktriangleright}d \rfloor$ to denote the number d. By default, we set $\lfloor d \rfloor = d$. For a vector $\bar{t} = (t_1, \ldots, t_m) \in \mathbb{B}^m$, we write $\lfloor \bar{t} \rfloor = (\lfloor t_1 \rfloor, \ldots, \lfloor t_m \rfloor) \in \mathbb{N}^m$. For two vectors $\bar{t}_1, \bar{t}_2 \in \mathbb{B}^m$, we define the dot product $\bar{t}_1 \cdot \bar{t}_2$ as $\lfloor \bar{t}_1 \rfloor \cdot \lfloor \bar{t}_2 \rfloor$.

Lemma B.4 *Let $\bar{d} \in \mathbb{B}^m$ such that $\bar{d} \notin \mathbb{N}^m$ and it does not contain any zero entry. For every $\bar{N} \in \mathbb{N}^m$ such that $\bar{N} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + (\bar{d} \cdot \bar{1})$, there exists a $\bar{d}$-regular graph G of size $\bar{N}$, where* $\deg(G) \leq (\bar{d} \cdot \bar{1}) + 1$.

Proof. The proof is a direct application of Lemma B.2. Let $\bar{d} = (d_1, \ldots, d_m)$. There are two cases.

- $\bar{N} \cdot \bar{d}$ is an even number.
  Then by Lemma B.2, there exists a $\lfloor \bar{d} \rfloor$-regular graph G of size $\bar{N}$, which is also a $\bar{d}$-regular graph. Obviously, $\deg(G) \leq (\bar{d} \cdot \bar{1}) + 1$.

- $\bar{N} \cdot \bar{t}$ is an odd number.
  By Lemma B.2, there exists an almost $\lfloor \bar{d} \rfloor$-regular graph G of size $\bar{N}$. Let $V_1 \cup \cdots \cup V_m$ and u be the culprit vertex in G. Let i be an index such that $d_i \in {}^{\blacktriangleright}\mathbb{N}$. Let $v \in V_i$ be a vertex not adjacent to u. Such a vertex v exists since $\bar{N} \cdot \bar{1} \geq (\bar{d} \cdot \bar{1})^4 + (\bar{d} \cdot \bar{1})$. Adding the edge (u,v) into G, we obtain a $\bar{d}$-regular graph of size $\bar{N}$ and $\deg(G) \leq \bar{d} \cdot \bar{1} + 1$.

This completes the proof of our lemma. ■

From Lemma B.4 we immediately get the following lemma.

Theorem B.5 *For every $\bar{d} \in \mathbb{B}^m$, there exists a Presburger formula $\mathsf{REG}_{\bar{d}}(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. There exists a $\bar{d}$-regular graph of size $\bar{N}$ if and only if $\mathsf{REG}_{\bar{d}}(\bar{N})$ holds.*

Proof. Apply Theorem B.3, if $\bar{d} \in \mathbb{N}^m$. Otherwise, we apply Lemma B.4 and run through the same argument as in the proof of Theorem B.3. We omit the details of the proof. ■

Proof of Theorem 4.1. We essentially run through the same argument in Subsection B.3 with a slight adjustment in the constants.

Given a matrix $D = \begin{pmatrix} \bar{t}_1 \\ \vdots \\ \bar{t}_\ell \end{pmatrix} \in \mathbb{N}^{\ell \times m}$ without zero column, we define the following set

$$I_D \quad := \quad \{\bar{N} \mid \bar{N} \cdot \bar{1} < 2\ell(D \cdot \bar{1} + \ell)^2 + 3\ell \text{ and there exists a } C\text{-regular graph of size } \bar{N}\}$$

Then, we define the formula $\mathsf{REG}_D(\bar{X})$ inductively, where $\bar{X} = (X_1, \ldots, X_m)$ as follows. For simplicity, we assume that D does not contain zero column. (If column i in D is a zero column, we simply add $X_i \geq 0$ and ignore the column.)

- When $\ell = 1$:

$$\mathsf{REG}_D(\bar{X}) \quad := \quad \mathsf{REG}_{\bar{t}_1}(\bar{X})$$

- When $\ell \geq 2$:

$$\begin{aligned} \mathsf{REG}_D(\bar{X}) \quad := \quad & \bigvee_{\bar{N} \in I_D} \bar{X} = \bar{N} \quad \vee \\ & \bigvee_{1 \leq j \leq \ell} \Big(\bar{X} \cdot \chi(\bar{t}_j) \geq 2(D \cdot \bar{1} + \ell)^2 + 3 \quad \wedge \quad \mathsf{REG}_{D - \bar{t}_j}(\bar{X}) \quad \wedge \quad \mathsf{REG}_{\bar{d}_j}(\bar{X}) \end{aligned}$$

  where $\chi(\bar{d}_j)$ is the characteristic vector of $\bar{d}_j$[††] and $D - \bar{d}_j$ denotes the matrix D without row j.

We can show by induction on ℓ that the following holds.

- If there exists a D-regular graph of size $\bar{N}$ then $\mathsf{REG}_D(\bar{N})$ holds.
- If $\mathsf{REG}_D(\bar{N})$ holds, then there exists a D-regular graph of size $\bar{N}$ with $\deg(G) \leq (D \cdot \bar{1}) + \ell$.

The details are similar to the one in the previous subsection, hence omitted.

C Proof of Theorem 4.2

In this section we present our proof of Theorem 4.2. We start with the following lemma which essentially states that if there exists a D-regular-complete graph of "big enough" size, then the matrix D itself has certain property.

Lemma C.1 *Let $G = (V, E_1, \ldots, E_\ell)$ be an ℓ-type D-regular graph, where $D \in \mathbb{B}^{\ell \times m}$ and $V = V_1 \cup \cdots \cup V_m$ is the partition of the regularity. Suppose that for each i, we have*

$$|V_i| \quad \geq \quad (D \cdot \bar{1}) + 1.$$

If G is also a complete graph, then for every $i, j \in \{1, \ldots, m\}$, there exists $l \in \{1, \ldots, \ell\}$ such that both $D_{l,i}, D_{l,j} \in {}^{\blacktriangleright}\mathbb{N}$.

[††]The characteristic vector of $\bar{c} = (c_1, \ldots, c_m) \in \mathbb{B}^m$ is defined as $\chi(c_1, \ldots, c_m) := (b_1, \ldots, b_m) \in \{0,1\}^m$ where $b_i = 0$ if $c_i = 0$, and $b_i = 1$ if $c_i \neq 0$.

Proof. Let $G = (V, E_1, \ldots, E_\ell)$ be a D-regular-complete graph where $D \in \mathbb{B}^{\ell \times m}$ and $V = V_1 \cup \cdots V_m$ the partition of the D-regularity Suppose that $\bar{t}_1, \ldots, \bar{t}_\ell$ are the row vectors of D and $|V_1|, \ldots, |V_m| \geq D \cdot \bar{1} + 1$.

For the sake of contradiction, we assume that that there exist $i, j \in \{1, \ldots, m\}$ such that for all $l \in \{1, \ldots, \ell\}$, either $D_{l,i} \in \mathbb{N}$ or $D_{l,j} \in \mathbb{N}$. We assume that $i \neq j$. The case when $i = j$ can be settled in a similar manner. We are going to count the number edges crossing between V_i and V_j and show that such number is strictly less than $|V_i| \cdot |V_j|$, which contradicts the fact that G is a complete graph.

For each $l \in \{1, \ldots, \ell\}$, we define the integer K_l as follows.

$$K_l \quad := \quad \begin{cases} |V_i| \cdot D_{l,i} & \text{if } D_{l,i} \in \mathbb{N} \\ |V_j| \cdot D_{l,j} & \text{if } D_{l,j} \in \mathbb{N} \end{cases}$$

If both $D_{l,i}$ and $D_{l,j}$ belong to $\mathbb{N}$, then we choose $|V_i| \cdot t_{l,i}$ for K_l. Then, we have

$$\sum_{1 \leq l \leq k} K_l \quad = \quad |V_i| \cdot \sum_{l \in L} D_{l,i} \quad + \quad |V_j| \cdot \sum_{l \notin L} D_{l,j}$$

where $L = \{l \mid D_{l,i} \in \mathbb{N}\}$. Now for each $l \in \{1, \ldots, k\}$, we define A_l as the set of E_l-edges crossing between V_i and V_j. Since G is a complete graph, we have

$$\sum_{1 \leq l \leq k} |A_l| \quad = \quad |V_i| \cdot |V_j|.$$

Moreover, G is D-regular, hence we have $|A_l| \leq K_l$ for each $l \in \{1, \ldots, k\}$. Thus,

$$\begin{aligned} \sum_{1 \leq l \leq k} |A_l| \quad &\leq \quad \sum_{1 \leq l \leq k} K_l \\ |V_i| \cdot |V_j| \quad &\leq \quad |V_i| \cdot \sum_{l \in L} D_{l,i} \quad + \quad |V_j| \cdot \sum_{l \notin L} D_{l,j}, \end{aligned}$$

However, such inequality cannot hold due to the fact that

$$\sum_{l \in L} D_{l,i} \quad + \quad \sum_{l \notin L} D_{l,j} \quad < \quad D \cdot \bar{1} + 1$$

and

$$|V_i|, |V_j| \quad \geq \quad D \cdot \bar{1} + 1.$$

This completes the proof of our lemma. ■

This lemma then motivates us to introduce the following notion of *easy* matrix. We way that a matrix $D \in \mathbb{B}^{\ell \times m}$ is an *easy* matrix, if for all $i, j \in \{1, \ldots, m\}$, there exists $l \in \{1, \ldots, \ell\}$ such that both $D_{l,i}, D_{l,j} \in {}^{\blacktriangleright}\mathbb{N}$. The following theorem shows the reason why it is called an easy matrix.

Lemma C.2 *Let $D \in \mathbb{B}^{\ell \times m}$ be an easy matrix. Then the following holds. There exists a D-regular-complete graph of size $\bar{N}$ if and only if $\mathsf{REG}_D(\bar{N})$ holds.*

Proof. Let $D \in \mathbb{B}^{\ell\times m}$ be an easy matrix. The "only if" direction follows directly from Theorem 4.1. We prove the "if" direction. Suppose $\mathsf{REG}_D(\bar{N})$ holds. By Theorem 4.1, there exists a D-regular graph $G = (V, E_1, \dots, E_\ell)$ of size $\bar{N}$. This graph G is not necessarily complete. So suppose $V = V_1 \cup \dots \cup V_m$ be the partition of the regularity. If G is not complete, then we perform the following. For all $u, v \in V$ such that $(u, v) \notin E_1 \cup \dots \cup E_\ell$, do the following.

- Let $u \in V_i$ and $v \in V_j$.
- Pick an index $l \in \{1, \dots, k\}$ such that $D_{l,i}, D_{l,j} \in {}^{\blacktriangleright}\mathbb{N}$.
  (Such an index l exists since D is an easy matrix.)
- Connect u and v with an E_l-edge.

The resulting graph is now complete and still D-regular. This completes the proof of our theorem. ■

The proof of Theorem 4.2 is essentially obtained by combining Lemmas C.1 and C.2. The main idea behind the formula $\mathsf{REG\text{-}COMP}_D(\bar{X})$ is as follows. Suppose D is not an easy matrix. Let J be the set of columns in D such that without those columns in J the matrix D is an easy matrix. Lemma C.1 tells us that if there exists a D-regular-complete graph of size $\bar{N}$, then the entries in $\bar{N}$ corresponding to the set J must be bounded from above. These entries then can be encoded as a Presburger formula via the notion of partial graph below.

A *partial graph* is a pair (P, f, n), where

- P is an ℓ-type graph;
- f is a function $V(P) \times \{E_1, \dots, E_\ell\} \times \{1, \dots, n\} \to \mathbb{B}$.

A completion of the partial graph (P, f, n) is an ℓ-type graph $G = (V, E_1, \dots, E_\ell)$ such that

- P is a subgraph of G, i.e., $V(P) \subseteq V(G)$ and $E_i(P) \subseteq E_i(G)$ for each $i = 1, \dots, \ell$;
- there is a partition of the vertices $V(G)\setminus V(P) = V_1 \cup \dots \cup V_n$, where for each $i = 1, \dots, \ell$, for each $j = 1, \dots, n$, for each vertex $v \in V(P)$ the number of vertices in V_j adjacent to v by E_i-edges is $f(v, E_i, j)$.

The gist of the proof of Theorem 4.2 is as follows. Given a partial graph (P, f, n) and an easy matrix $D \in \mathbb{B}^{\ell\times n}$, we are going to construct a formula $\Psi_{(P,f,n),D}$ such that there is a completion G of the partial graph (P, f, n) with an additional condition that the partition $V(G) \setminus V(P) = V_1 \cup \dots \cup V_n$ satisfies the following. For each $i = 1, \dots, \ell$, for each $j = 1, \dots, n$, for each vertex $v \in V_j$, $\deg_{E_i}(v) = D_{i,j}$. The construction of the formula $\Psi_{(P,f,n),D}$ can be done by induction on the number of vertices in P.

We present the details in the following paragraph. We need to define the operation addition $+$ on $\mathbb{B}$ as follows.

$$
\begin{aligned}
d_1 + d_2 &:= d_1 + d_2 \\
{}^{\blacktriangleright}d_1 + d_2 &:= {}^{\blacktriangleright}(d_1 + d_2) \\
d_1 + {}^{\blacktriangleright}d_2 &:= {}^{\blacktriangleright}(d_1 + d_2) \\
{}^{\blacktriangleright}d_1 + {}^{\blacktriangleright}d_2 &:= {}^{\blacktriangleright}(d_1 + d_2)
\end{aligned}
$$

Obviously, we can extend $+$ to the spaces $\mathbb{B}^m$ and $\mathbb{B}^{\ell\times m}$, where they are performed component-wise. We also need to define the subtraction on ${}^{\blacktriangleright}\mathbb{N}$ as follows. For $d_1, d_2 \in \mathbb{N}$,

$$
({}^{\blacktriangleright}d_1) - d_2 = \begin{cases} {}^{\blacktriangleright}(d_1 - d_2) & \text{if } d_1 \geq d_2 \\ {}^{\blacktriangleright}0 & \text{otherwise} \end{cases}
$$

For a matrix $D \in \mathbb{B}^{\ell \times m}$ and a set $I \subseteq \{1, \ldots, m\}$, we write D_I the matrix in $\mathbb{B}^{k \times |I|}$ obtained from D by deleting the columns not in I. For two matrices $D \in \mathbb{B}^{\ell \times m_1}$ and $D' \in \mathbb{B}^{\ell \times m_2}$, we write $(D|D')$ to denote the matrix in $\mathbb{B}^{\ell \times (m_1+m_2)}$, where the first m_1 columns are D, while the last m_2 columns are D'.

For a function $f : U \times \{E_1, \ldots, E_\ell\} \times \{1, \ldots, n\} \to \mathbb{B}$, an ℓ-expansion of the function f is a function

$$g \ : \ U \times \{E_1, \ldots, E_\ell\} \times (\{1, \ldots, n\} \times \{1, \ldots, \ell+1\}) \quad \to \quad \mathbb{B}$$

such that for each $u \in U$, $l \in \{1, \ldots, \ell+1\}$, $i \in \{1, \ldots, n\}$,

$$\sum_{1 \leq j \leq \ell+1} g(u, E_l, (i,j)) \quad = \quad f(u, E_l, i).$$

Abusing the notation, we can view the domain of g as $U \times \{E_1, \ldots, E_\ell\} \times \{1, \ldots, n(\ell+1)\}$, instead of $U \times \{E_1, \ldots, E_\ell\} \times (\{1, \ldots, n\} \times \{1, \ldots, \ell+1\})$. Also we write the function $f \setminus \{u\}$ obtained from the same function f but restricted to $V \setminus \{u\}$.

For a matrix $D \in \mathbb{B}^{\ell \times m}$, we define an operator $\xi(D) \in \mathbb{B}^{\ell \times (\ell m + m)}$ as follows. Let $D = (\bar{c}_1^t \mid \cdots \mid \bar{c}_m^t)$, that is, $\bar{c}_1^t, \ldots, \bar{c}_m^t$ are the column vectors of D.

$$\xi(D) \quad := \quad (D_1 \mid \bar{c}_1^t \mid D_2 \mid \bar{c}_2^t \mid \cdots \mid D_m \mid \bar{c}_m^t),$$

where $D_i = (\bar{c}_i^t \mid \cdots \mid \bar{c}_i^t) - I_\ell \in \mathbb{B}^{\ell \times \ell}$ and I_ℓ is the identity $(\ell \times \ell)$ matrix. We remark that if D is an easy matrix, then so is $\xi(D)$.

Lemma C.3 *Let (P, f, m) be a partial graph. For every matrix $D \in \mathbb{B}^{k \times m}$, there exists a Presburger formula $\Psi_{(P,f,m),D}(\bar{X})$, where $\bar{X} = (X_1, \ldots, X_m)$ such that the following holds. The sentence $\Psi_{(P,f,m),D}(\bar{N})$ holds if and only if there exists a graph $G = (V, E_1, \ldots, E_\ell)$ in which there is a partition $V = V_0 \cup V_1 \cup \cdots \cup V_m$ such that*

- *$V_0 = V(P)$;*
- *$E_l(P) \subseteq E_l$ for every $l \in \{1, \ldots, \ell\}$;*
- *$(|V_1|, \ldots, |V_m|) = \bar{N}$;*
- *for each $u \in V(P)$, for each $l \in \{1, \ldots, \ell\}$, for each $j \in \{1, \ldots, m\}$, the number of E_l-edges from u to V_j is $f(u, E_l, j)$;*
- *for each $j \in \{1, \ldots, m\}$, for each $v \in V_j$, for each $l \in \{1, \ldots, \ell\}$, the degree $\deg_{E_l}(v)$ in G is $D_{l,j}$.*

Proof. Let (P, f, m) be a partial graph, and let $\alpha_1, \ldots, \alpha_s$ be the enumeration of the vertices in P.

For an integer $h \in \{0, 1, \ldots, s\}$, a function $g : \{\alpha_1, \ldots, \alpha_h\} \times \{E_1, \ldots, E_\ell\} \times \{1, \ldots, n\}$ and an easy matrix $C \in \mathbb{B}^{\ell \times n}$, we define the formula $\Phi_{h,g,C}$ inductively on h as follows.

$$\begin{aligned}
\Phi_{0,\emptyset,n,C}(\bar{X}) \quad &:= \quad \mathsf{REG}_C(\bar{X}) \\
\Phi_{h+1,g,n,C}(\bar{X}) \quad &:= \quad \exists Z_{1,1} \cdots \exists Z_{1,\ell+1}\ \exists Z_{2,1} \cdots \exists Z_{2,\ell+1} \cdots\cdots \exists Z_{n,1} \cdots \exists Z_{n,\ell+1} \\
&\qquad \bigwedge_{1 \leq i \leq n} Z_{i,1} + \cdots + Z_{i,\ell+1} = X_i \\
&\quad \wedge \bigwedge_{j \text{ s.t. } g(\alpha_{h+1}, E_j, i) \in \mathbb{N}} Z_{i,j} = g(\alpha_{h+1}, E_j, i)
\end{aligned}$$

$$\wedge \bigwedge_{j \text{ s.t. } g(\alpha_{h+1},E_j,i)\in \blacktriangleright\mathbb{N}} Z_{i,j} \geq g(\alpha_{h+1}, E_j, i)$$

$$\wedge \bigvee_{g' \text{ is an } \ell\text{-extension of } g} \Phi_{h,g'\setminus\{\alpha_{h+1}\},\xi(C)}((Z_{1,1}, \ldots, Z_{1,\ell+1}), \ldots, (Z_{n,1}, \ldots, Z_{n,\ell+1}))$$

We note that when $h = 0$, then the function g is the empty function. Now the desired formula $\Psi_{(P,f,m),D}$ is defined as follows.

$$\Psi_{(P,f,m),D}(\bar{X}) \quad := \quad \Phi_{s,f,m,D}(\bar{X}).$$

That the formula $\Psi_{(P,f,m),D}(\bar{X})$ is the desired formula comes from the following claim. For $h \in \{0, 1, \ldots, s\}$, we are going to write P_h as the subgraph of P restricted to the vertices $\alpha_1, \ldots, \alpha_h$. By default, we set P_0 to be the empty graph.

Claim 1 *For an integer $h \in \{0, 1, \ldots, s\}$, a function $g : \{\alpha_1, \ldots, \alpha_h\} \times \{E_1, \ldots, E_\ell\} \times \{1, \ldots, n\}$ and a matrix $C \in \mathbb{B}^{\ell\times n}$, the following holds. The sentence $\Phi_{h,g,C}(\bar{N})$ holds if and only if there exists a graph $G = (V, E_1, \ldots, E_\ell)$ in which there is a partition $V = V_0 \cup V_1 \cup \cdots \cup V_m$, such that*

- *$V_0 = V(P_h)$;*
- *$E_l(P_h) \subseteq E_l$ for every $l \in \{1, \ldots, \ell\}$;*
- *$(|V_1|, \ldots, |V_m|) = \bar{N}$;*
- *for each $\alpha \in V(P_h)$, for each $l \in \{1, \ldots, \ell\}$, for each $j \in \{1, \ldots, m\}$, the number of E_l-edges from α to V_j is $g(\alpha, E_l, j)$;*
- *for each $j \in \{1, \ldots, m\}$, for each $v \in V_j$, for each $l \in \{1, \ldots, \ell\}$, the degree $\deg_{E_l}(v)$ in G is $D_{l,j}$.*

We are going to prove the claim by induction on the integer h. The base case $h = 0$ holds due to Lemma C.2. For the induction hypothesis, we assume that the claim holds for case h.

We are going to prove the claim for case $h + 1$ for the induction step. First we prove the "if" direction. Suppose there exists a graph $G = (V, E_1, \ldots, E_\ell)$ in which there is a partition $V_0 \cup V_1 \cup \cdots \cup V_m$ of V, such that

- $V_0 = V(P_{h+1})$;
- $E_l(P_{h+1}) \subseteq E_l$ for every $l \in \{1, \ldots, \ell\}$;
- $(|V_1|, \ldots, |V_n|) = \bar{N}$;
- for each $\alpha \in V(P_{h+1})$, for each $l \in \{1, \ldots, \ell\}$, for each $j \in \{1, \ldots, n\}$, the number of E_l-edges from α to V_j is $g(\alpha, E_l, j)$;
- for each $j \in \{1, \ldots, n\}$, for each $v \in V_j$, for each $l \in \{1, \ldots, \ell\}$, the degree $\deg_{E_l}(v)$ in G is $C_{l,j}$;

Now for each $j \in \{1, \ldots, n\}$, we further partition V_j into $V_{j,1} \cup \cdots \cup V_{j,\ell} \cup V_{j,\ell+1}$, where $V_{j,l}$ is the set of vertices in V_j connected to α_{h+1} by E_l-edges for $l = 1, \ldots, \ell$ and $V_{j,\ell+1}$ is the set of vertices in V_j that is not connected to α_{h+1}. It is straightforward application of the induction hypothesis that the sentence $\Phi_{h+1,g,C}(\bar{N})$ holds.

The proof for the "only if" direction is very similar. For completeness, we prove it here. Suppose $\Phi_{h+1,g,C}(\bar{N})$ holds. Let $M_{1,1}, \ldots, M_{1,\ell+1}, M_{2,1}, \ldots, M_{2,\ell+1}, \ldots, M_{n,1}, \ldots, M_{n,\ell+1}$ be

the witnesses to the variables $Z_{1,1},\dots,Z_{1,\ell+1},Z_{2,1},\dots,Z_{2,\ell+1},\dots,Z_{n,1},\dots,Z_{n,\ell+1}$ such that the sentence $\Phi_{h+1,g,C}(\bar{N})$ holds. By the definition of $\Phi_{h+1,g,C}(\bar{N})$, there exists a function g' which is a ℓ-extension of g such that

$$\Phi_{h,g'\setminus\{\alpha_{h+1}\},\xi(C)}((M_{1,1},\dots,M_{1,\ell}),\dots,(M_{n,1},\dots,M_{n,\ell})) \qquad \text{holds.}$$

By the induction hypothesis, there exists a graph $G=(V,E_1,\dots,E_\ell)$ in which there is a partition $V_0\cup V_{1,1}\cup\dots\cup V_{n,\ell+1}$ of V, such that

- $V_0=V(P_h)$;
- $E_l(P_h)\subseteq E_l$ for every $l\in\{1,\dots,\ell\}$;
- $((|V_{1,1}|,\dots,|V_{1,\ell+1}|),\dots,(|V_{n,1}|,\dots,|V_{n,\ell+1}|))=((M_{1,1},\dots,M_{1,\ell+1}),\dots,(M_{n,1},\dots,M_{n,\ell+1}))$;
- for each $\alpha\in V(P_h)$, for each $l\in\{1,\dots,\ell\}$, for each $(j,j')\in\{1,\dots,n\}\times\{1,\dots,\ell\}$, the number of E_l-edges from α to $V_{j,j'}$ is $g'(\alpha,E_l,(j,j'))$;
- for each $(j,j')\in\{1,\dots,n\}\times\{1,\dots,\ell\}$, for each $v\in V_{j,j'}$, for each $l\in\{1,\dots,\ell\}$, the degree $\deg_{E_l}(v)$ in G is $C_{l,j,j'}$.

Now for each $i\in\{1,\dots,n\}$, we define $V_i=V_{i,1}\cup\dots\cup V_{i,\ell+1}$, and the set of E_l-edges E'_l as E_l plus connecting the vertices in $V_{i,l}$ to the vertex α_{h+1} with E_l edges. By the constraint

$$\begin{aligned} &\bigwedge_{1\le i\le n} Z_{i,1}+\dots+Z_{i,\ell+1}=X_i\\ \wedge &\bigwedge_{j \text{ s.t. } g(\alpha_{h+1},E_j,i)\in\mathbb{N}} Z_{i,j}=g(\alpha_{h+1},E_j,i)\\ \wedge &\bigwedge_{j \text{ s.t. } g(\alpha_{h+1},E_j,i)\in^{\blacktriangleright}\mathbb{N}} Z_{i,j}\ge g(\alpha_{h+1},E_j,i)\end{aligned}$$

we get that the number of E_l-edges from α_{h+1} to V_i is $g(\alpha_{h+1},E_l,i)$. Hence, this way a partition $V'_0\cup V_1\cup\dots\cup V_m$ of V, such that

- $V'_0=V(P_{h+1})$;
- $(|V_1|,\dots,|V_m|)=\bar{N}$;
- for each $\alpha\in V(P_h)$, for each $l\in\{1,\dots,\ell\}$, for each $j\in\{1,\dots,m\}$, the number of E_l-edges from α to V_j is $g(\alpha,E_l,j)$;
- for each $j\in\{1,\dots,m\}$, for each $v\in V_j$, for each $l\in\{1,\dots,\ell\}$, the degree $\deg_{E_l}(v)$ in G is $D_{l,j}$.

This completes the proof of our claim, and hence, our lemma. ■

Note that in Lemma C.3 above the completion graph G is not necessarily complete. To get a complete graph G, we have to insist that P itself is a complete graph and the matrix D is easy as stated below.

Lemma C.4 *Let (P,f,m) be a partial graph and P itself is a complete graph. For every easy matrix $D\in\mathbb{B}^{k\times m}$, there exists a Presburger formula $\Psi_{(P,f,m),D}(\bar{X})$, where $\bar{X}=(X_1,\dots,X_m)$ such that the following holds. The sentence $\Psi_{(P,f,m),D}(\bar{N})$ holds if and only if there exists a* complete *graph $G=(V,E_1,\dots,E_\ell)$ in which there is a partition $V=V_0\cup V_1\cup\dots\cup V_m$ such that*

- $V_0 = V(P)$;
- $E_l(P) \subseteq E_l$ *for every* $l \in \{1, \ldots, \ell\}$;
- $(|V_1|, \ldots, |V_m|) = \bar{N}$;
- *for each* $u \in V(P)$, *for each* $l \in \{1, \ldots, \ell\}$, *for each* $j \in \{1, \ldots, m\}$, *the number of* E_l*-edges from* u *to* V_j *is* $f(u, E_l, j)$;
- *for each* $j \in \{1, \ldots, m\}$, *for each* $v \in V_j$, *for each* $l \in \{1, \ldots, \ell\}$, *the degree* $\deg_{E_l}(v)$ *in* G *is* $D_{l,j}$.

Proof. The proof is very similar as above. We simply add the constraint $Z_{i,\ell+1} = 0$ for each $i = 1, \ldots, n$ in the definition of the formula $\Phi_{h+1,g,n,C}(\bar{X})$ above. ■

Proof of Theorem 4.2. The proof is a direct application of Lemma C.3. Let $D \in \mathbb{B}^{\ell \times m}$, not necessarily easy matrix. For a subset $J \subseteq \{1, \ldots, m\}$, we write D_J as the matrix D with the columns only from J.

A partial graph (P, f, n) is compatible with (D, J), if

- $|J| = n$;
- there is a partition $V(P) = V_{j_1} \cup \cdots \cup V_{j_{m-n}}$, where $\{j_1, \ldots, j_{m-n}\} = \{1, \ldots, m\} - J$ and $|V_{j_i}| \le D \cdot \bar{1}$ for each j_i;
- for each $1 \le i \le m-n$, for each $v \in V_{j_i}$, for each E_l, we have $\sum_{1 \le k \le m} f(v, E_l, k) = D_{l,j_i}$.

Now the formula $\mathsf{REG\text{-}COMP}_D(\bar{X})$ can be defined as follows.

$$\bigvee_{J=\{j_1,\ldots,j_m\} \text{ s.t. } D_J \text{ is easy}} \quad \bigvee_{(P,f,m) \text{ s.t. it is compatible with } (D,J) \text{ and we index the partition } V(P)=\bigcup_{i \notin J} V_i}$$
$$\Psi_{P,f,m,D_J}(X_{j_1}, \ldots, X_{j_m}) \quad \wedge \quad \bigwedge_{i \notin J} X_i = |V_i|$$

The correctness of the formula follows immediately from Lemma C.4.

D Proof of Theorem 5.1

For convenience, we restate Theorem 5.1 here.

Theorem 5.1. *Let* $C \in \mathbb{B}^{\ell \times m}$ *and* $D \in \mathbb{B}^{\ell \times n}$. *There is a Presburger formula* $\mathsf{BiREG}_{C,D}(\bar{X}, \bar{Y})$, *where* $\bar{X} = (X_1, \ldots, X_m)$ *and* $\bar{Y} = (Y_1, \ldots, Y_n)$ *such that the following holds. There exists an* ℓ*-type* (C, D)*-biregular graph of size* $(\bar{M}, \bar{N})$ *if and only if* $\mathsf{BiREG}_{C,D}(\bar{M}, \bar{N})$ *holds.*

The structure of the proof is the same as in Appendix B. The essence of the theorem lies on the case when $C, D \in \mathbb{N}^{\ell \times m}$. This is what we will consider first for the sake of presentation. The proof is divided into three successive steps.

(1) First, we consider the simplest case when $C, D \in \mathbb{N}^{1 \times 1}$ in Subsection D.1.

(2) Then we generalise it to the case $C \in \mathbb{N}^{1 \times m}$ and $D \in \mathbb{N}^{1 \times n}$ in Subsection D.2.

(3) Finally we consider the case $C \in \mathbb{N}^{\ell \times m}$ and $D \in \mathbb{N}^{\ell \times n}$ in Subsection D.3.

When C, D contain an element of $\mathbb{B}$, it will be easier since there is more freedom in adding more edges. However, the proof will be more tedious. For this reason, we postpone the case when D contains an element of $\mathbb{B}$ until Subsection D.4.

D.1 When $C, D \in \mathbb{N}^{1\times 1}$

In this subsection we deal with the simplest case of (C, D)-biregular graph: the case when C and D consist of only one entry c and d, respectively. For convenience, we will simply write (c, d)-biregular graphs, instead of a more cumbersome notation $((c), (d))$-biregular graph.

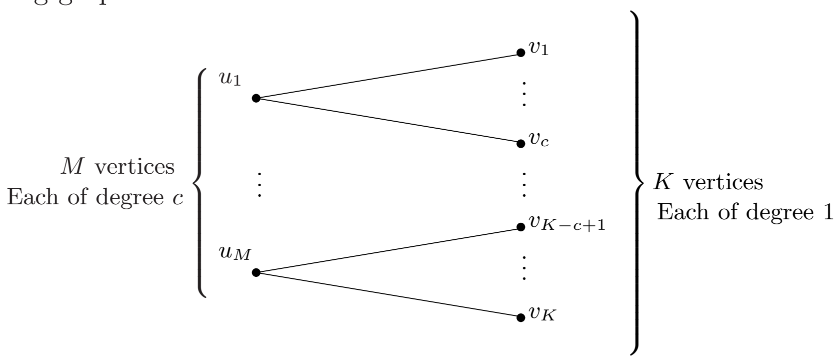
In this subsection and the next, we are dealing only with 1-type graph. Hence, we are going to write $\deg(u)$ to denote the degree of the vertex u, instead of $\deg_E(u)$.

We start with the following proposition, which again appears to be a folklore in the graph theory. We represent the proof here since we are going to need it.

Proposition D.1 *Let $c, d \geq 0$. For every $M, N \in \mathbb{N}$, the following holds. There exists a (c, d)-biregular graph of size (M, N) if and only if $N \geq c$, $M \geq d$ and $M \cdot c = N \cdot d$.*

Proof. Let $c, d \geq 0$, and let $M, N \in \mathbb{N}$. The "only if" directions is straightforward.

We first prove the "if" direction. Let $M \geq d$, $N \geq c$. Let $K = Mc = Nd$. First, we construct the following graph.



On the left side, we have M vertices, and each has degree c. On the right side, we have $K = Nd$ vertices, and each has degree 1. We are going to merge every d vertices on the right side into one vertex of degree d. The merging is as follows. We merge every d vertices $v_i, v_{i+N}, \ldots, v_{i+(d-1)N}$ into one vertex for every $i = 1, \ldots, N$. Since $K = Nd$, it is possible to do such merging. Moreover, $N \geq c$, hence we do not have multiple edges between two vertices. Thus, we obtain the desired (c, d)-biregular graph of size (M, N). This completes the proof of Proposition D.1. ■

D.2 When $C \in \mathbb{N}^{1\times m}$ and $D \in \mathbb{N}^{1\times n}$

In this subsection we consider the case when C and D consist of only one vector each. In this case, we are going to write $(\bar{c}, \bar{d})$-biregular graph, where $\bar{c}$ and $\bar{d}$ are the only vectors of C and D, respectively.

Lemma D.2 *Let $\bar{c} \in \mathbb{N}^m$ and $\bar{d} \in \mathbb{N}^n$ and both do not contain zero entry. For each $\bar{M} \in \mathbb{N}^m$ and $\bar{N} \in \mathbb{N}^n$ such that*

$$\bar{M} \cdot \bar{1} + \bar{N} \cdot \bar{1} \quad \geq \quad 2(\bar{c} \cdot \bar{1})(\bar{d} \cdot \bar{1}) + 3,$$

the following holds. There exists a $(\bar{c}, \bar{d})$-biregular graph of size $(\bar{M}, \bar{N})$ if and only if $\bar{M} \cdot \bar{c} = \bar{N} \cdot \bar{d}$.

Proof. Let $\bar{c} \in \mathbb{N}^m$ and $\bar{d} \in \mathbb{N}^n$, both do not contain zero entry. Let $\bar{M} \in \mathbb{N}^m$, $\bar{N} \in \mathbb{N}^n$ such that

$$\bar{M} \cdot \bar{1} + \bar{N} \cdot \bar{1} \quad \geq \quad 2(\bar{c} \cdot \bar{1})(\bar{d} \cdot \bar{1}) + 3.$$

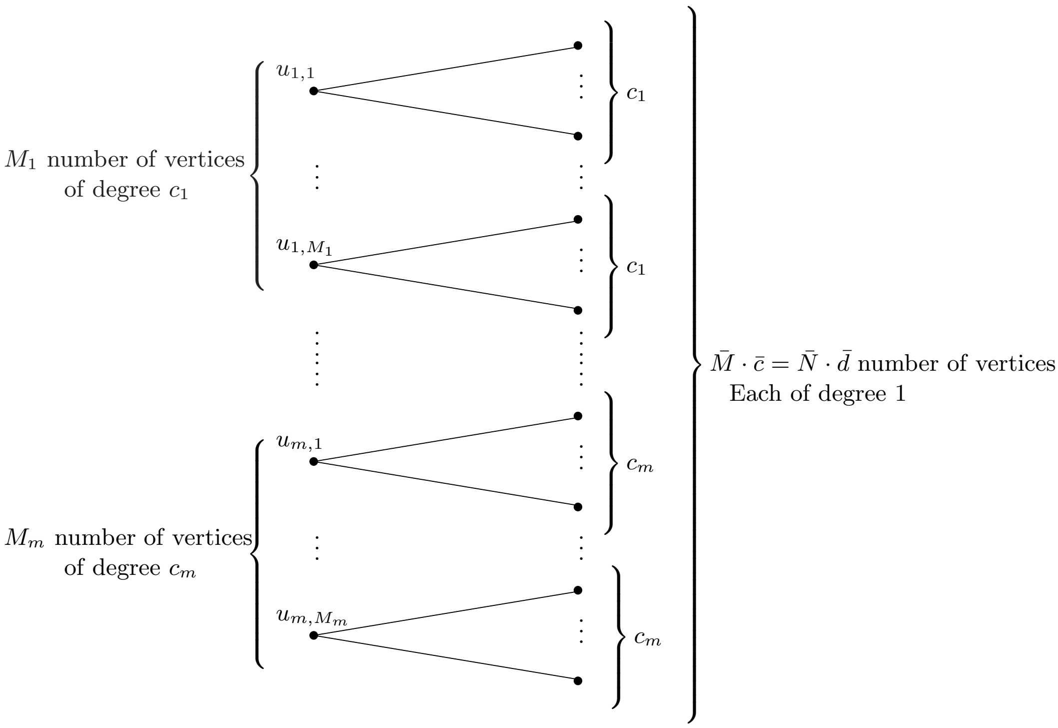
The "only if" direction is straightforward. If G is a $(\bar{c}, \bar{d})$-biregular graph of size $(\bar{M}, \bar{N})$, then the number of edges in G is precisely $\bar{M} \cdot \bar{c} = \bar{N} \cdot \bar{d}$.

Now we prove the "if" part. Suppose $\bar{M} \in \mathbb{N}^m$, $\bar{N} \in \mathbb{N}^n$ such that $\bar{M} \cdot \bar{c} = \bar{N} \cdot \bar{d}$. Let $\bar{c} = (c_1, \ldots, c_m)$ and $\bar{d} = (d_1, \ldots, d_n)$, and $\bar{M} = (M_1, \ldots, M_m)$ and $\bar{N} = (N_1, \ldots, N_n)$.

We are going to construct a $(\bar{c}, \bar{d})$-biregular graph of size $(\bar{M}, \bar{N})$. We first construct a bipartite graph, in which

- the left side has $\bar{M} \cdot \bar{1}$ vertices, where there are M_1 vertices $u_{1,1}, \ldots, u_{1,M_1}$ and each has c_1 vertices adjacent to it; there are M_2 vertices $u_{2,1}, \ldots, u_{2,M_2}$, and each has c_2 vertices adjacent to it; etc;
- the right side has $\bar{M} \cdot \bar{c}$ number of vertices, each of degree one.

See the illustration below.



Obviously on the left side there are M_1 vertices of degree c_1, M_2 nodes of degree c_2, etc. However, on the right side there are $\bar{M} \cdot \bar{c}$ vertices and each is of degree one. We are going to do some merging of the vertices on the right side so that there are exactly N_1 vertices of degree d_1, N_2 vertices of degree d_2, etc. We do the following.

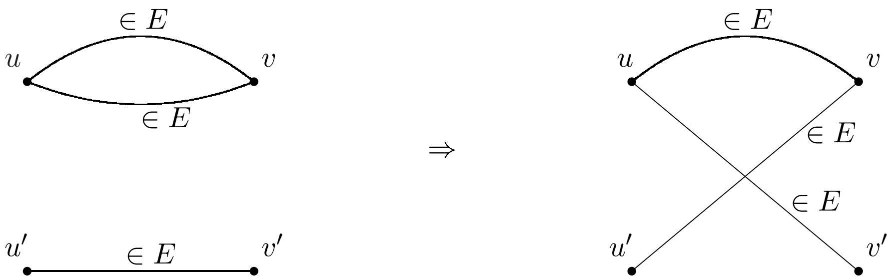
We "group" the vertices on the right side into $V_1, \ldots, V_n$ where V_1 has $N_1 d_1$ vertices, V_2 has $N_2 d_2$ vertices, etc. Such grouping is possible because $\bar{M} \cdot \bar{c} = \bar{N} \cdot \bar{d}$.

For each $i \in \{1, \ldots, n\}$, we do the following. We merge d_i vertices in V_i into one vertex, so that each vertex in V_i has degree d_i. The merging is done like in the proof of Proposition D.1. Let $V_i = \{v_{i,1}, \ldots, v_{i,K_i}\}$ where $K_i = N_i d_i$. We merge the vertices

$v_1, v_{N_i+1}, v_{2N_i+1}, \dots, v_{(d_i-1)N_i+1}$ into one vertex; the vertices $v_2, v_{N_i+2}, v_{2N_i+2}, \dots, v_{(d_i-1)N_i+2}$ into one vertex; and so on.

After such merging, each vertex in V_i has degree d_i. However, it is possible that after we do the merging, we have "parallel" edges, i.e. more than one edges between two vertices. (See the left side of the illustration below.) We are going to "remove" such parallel edges one by one until there is no more parallel edges. The trick is similar as the one in Subsection B.3.

Suppose we have parallel edges between the vertices u and v. We pick an edge (u', v') such that u' is not adjacent to v and v' is not adjacent to u. (See the left side of the illustration below.)



Such an edge (u', v') exists since the number of vertices reachable in distance 2 from the vertices u and v is $\leq 2(\bar{c}\cdot\bar{1})(\bar{d}\cdot\bar{1})+2$ and the number of vertices is $\bar{M}\cdot\bar{1}+\bar{N}\cdot\bar{1} \geq 2(\bar{c}\cdot\bar{1})(\bar{d}\cdot\bar{1})+3$.

Now we delete the edges (u', v') and one of the parallel edge (u, v), replace it with the edges (u, v') and (u', v), as illustrated on the right side of the illustration above. We perform such operation until there is no more parallel edges. This completes the proof. ■

The following theorem is a straightforward application of Lemma D.2.

Theorem D.3 *For every $\bar{c} \in \mathbb{N}^m$ and $\bar{d} \in \mathbb{N}^n$, there exists a Presburger formula $\mathsf{BiREG}_{(\bar{c},\bar{d})}(\bar{X}, \bar{Y})$, where $\bar{X} = (X_1, \dots, X_m)$ and $\bar{Y} = (Y_1, \dots, Y_n)$ such that the following holds. There exists a $(\bar{c}, \bar{d})$-biregular graph of size $(\bar{M}, \bar{N})$ if and only if the sentence $\mathsf{BiREG}_{(\bar{c},\bar{d})}(\bar{M}, \bar{N})$ holds.*

Proof. The proof is a direct application of Lemma D.2. It can be proved with the same line of reasoning as Theorem B.3. We omit the details. ■

D.3 When $C \in \mathbb{N}^{\ell\times m}$ and $D \in \mathbb{N}^{\ell\times n}$

In this subsection we are going to generalise Theorem D.3 to the case $\ell \geq 1$.

Theorem D.4 *For every $C \in \mathbb{N}^{\ell\times m}$ and $D \in \mathbb{N}^{\ell\times n}$, there exists a Presburger formula $\mathsf{BiREG}_{C,D}(\bar{X}, \bar{Y})$, where $\bar{X} = (X_1, \dots, X_m)$ and $\bar{Y} = (Y_1, \dots, Y_n)$ such that the following holds. There exists a (C, D)-biregular ℓ-type graph of size $(\bar{M}, \bar{N})$ if and only if the sentence $\mathsf{BiREG}_{C,D}(\bar{M}, \bar{N})$ holds.*

Proof. The strategy of the proof here is also very similar to the one in Theorem 4.1. In fact, we repeat the same line of reasoning with minor adjustment to the bipartite graphs setting.

Given $C = \begin{pmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_\ell \end{pmatrix} \in \mathbb{N}^{\ell \times m}$ and $D = \begin{pmatrix} \bar{d}_1 \\ \vdots \\ \bar{d}_\ell \end{pmatrix} \in \mathbb{N}^{\ell \times n}$ without zero column. We define the following set

$$I_{C,D} \quad := \quad \left\{ (\bar{M}, \bar{N}) \;\middle|\; \begin{array}{l} \bar{M} \cdot \bar{1} + \bar{N} \cdot \bar{1} < 2\ell(C \cdot \bar{1})(D \cdot \bar{1}) + 3\ell \text{ and} \\ \text{there exists a } (C, D)\text{-biregular graph of size } (\bar{M}, \bar{N}) \end{array} \right\}$$

Then, we define the formula $\mathsf{BiREG}_{C,D}(\bar{X}, \bar{Y})$ inductively, where $\bar{X} = (X_1, \ldots, X_m)$ and $\bar{Y} = (Y_1, \ldots, Y_n)$ as follows. For simplicity, we assume that C and D do not contain zero column. (If column i in matrix C (or, D, respectively) is zero column, then we add the constraint $X_i \geq 0$ (or, $Y_i \geq 0$, respectively) and ignore that column.)

- When $\ell = 1$:

$$\mathsf{BiREG}_{C,D}(\bar{X}, \bar{Y}) \quad := \quad \mathsf{BiREG}_{\bar{c}_1, \bar{d}_1}(\bar{X}, \bar{Y})$$

- When $\ell \geq 2$:

$$\begin{aligned} \mathsf{BiREG}_{C,D}(\bar{X}, \bar{Y}) \quad := \quad & \bigvee_{(\bar{M},\bar{N}) \in I_{C,D}} \bar{X} = \bar{M} \ \wedge \ \bar{Y} = \bar{N} \\ & \vee \bigvee_{1 \leq j \leq \ell} \Big(\bar{X} \cdot \chi(\bar{c}_j) + \bar{Y} \cdot \chi(\bar{d}_j) \geq 2(C \cdot \bar{1})(D \cdot \bar{1}) + 3 \\ & \qquad\qquad \wedge \quad \mathsf{REG}_{C - \bar{c}_j, D - \bar{d}_j}(\bar{X}, \bar{Y}) \quad \wedge \quad \mathsf{REG}_{\bar{c}_j, \bar{d}_j}(\bar{X}, \bar{Y}) \Big) \end{aligned}$$

  where $\chi(\bar{c}_j), \chi(\bar{d}_j)$ are the characteristic vectors of $\bar{c}_j$ and $\bar{d}_j$ defined in Subsection B.3 and $C - \bar{c}_j$, $D - \bar{d}_j$ denote the matrices C without row j and D without row j, respectively.

The correctness of our formula can be proved in a similar manner as Theorem 4.1, hence we omit the details. This completes our proof. ■

D.4 When C and D contain elements from $\mathbb{B}$

In this subsection we are going prove Theorem D.4 when C, D contain elements from $\mathbb{B}$. As warm up, we consider the following easy case, from which Theorem 5.2 can be easily generalised.

Proposition D.5 *Let $c, d \geq 0$. For every $M, N \in \mathbb{N}$, the following holds.*

(a) There exists a $(c, {}^{\blacktriangleright}d)$-biregular graph of size (M, N) if and only if $M \geq d$, $N \geq c$ and $Mc \geq Nd$.

(b) There exists a $({}^{\blacktriangleright}c, {}^{\blacktriangleright}d)$-biregular graph of size (M, N) if and only if $M \geq d$, $N \geq c$.

Proof. Let $c, d \geq 0$, and let $M, N \in \mathbb{N}$. The "only if" directions on all parts (a) and (b) are straightforward.

The "if" direction of part (a) can be proved in a similar manner as Proposition D.1. First, we construct a $(c, 1)$-bipartite graph of size (M, Mc). Since $Mc \geq Nd$, when we do the merging

among the vertices on the right side, we merge at least d vertices $v_i, v_{i+N}, \ldots, v_{i+(d-1)N}$ into one vertex for every $i = 1, \ldots, N$. Hence, after the merging each vertex has degree $\geq d$.

The "if" direction of part (b) is straightforward. If $Mc \geq Nd$, then we construct a $(c, {\blacktriangleright}d)$-biregular graph of size (M, N). If $Mc < Nd$, then we construct a $({\blacktriangleright}c, d)$-biregular graph of size (M, N). This completes the proof of Proposition D.5. ∎

Proposition D.5 can be generalised to the case when C and D consist of one row, as stated in the following theorem.

Theorem D.6 *For every $\bar{c} \in \mathbb{B}^m$, $\bar{d} \in \mathbb{B}^n$, there exists a Presburger formula $\mathsf{BiREG}_{(\bar{c},\bar{d})}(\bar{X}, \bar{Y})$ such that the following holds. There exists a $(\bar{c}, \bar{d})$-biregular graph of size $(\bar{M}, \bar{N})$ if and only if the statement $\mathsf{BiREG}_{(\bar{c},\bar{d})}(\bar{M}, \bar{N})$ holds.*

We postpone the proof of Theorem D.6 until later. It basically follows directly from the following lemma.

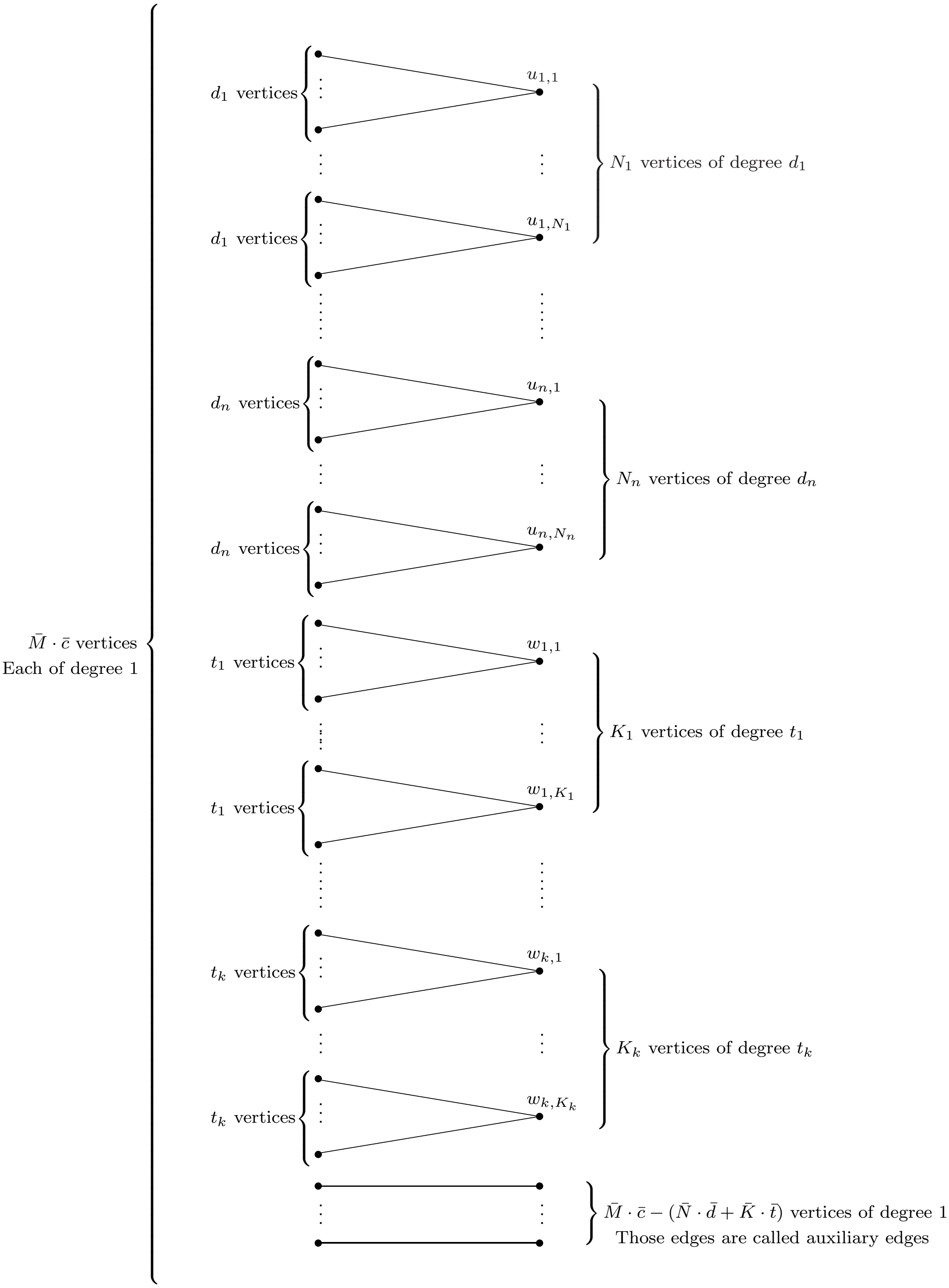
Lemma D.7 *Let $\bar{c} \in \mathbb{N}^m$ and $\bar{d} \in \mathbb{N}^n$ and $\bar{t} \in ({\blacktriangleright}\mathbb{N})^k$. For all $\bar{M} \in \mathbb{N}^m$ and $\bar{N} \in \mathbb{N}^n$ and $\bar{K} \in \mathbb{N}^k$ such that $\bar{M} \cdot \bar{1} + \bar{N} \cdot \bar{1} + \bar{K} \cdot \bar{1} \geq 2(\bar{c} \cdot \bar{1})(\bar{d} \cdot \bar{1} + \bar{t} \cdot \bar{1}) + 3$ and $\bar{K} \cdot \bar{1} \geq \bar{c} \cdot \bar{1}$, the following holds. There exists a $(\bar{c}, (\bar{d}, \bar{t}))$-biregular graph G of size $(\bar{M}, (\bar{N}, \bar{K}))$ if and only if $\bar{M} \cdot \bar{c} \geq \bar{N} \cdot \bar{d} + \bar{K} \cdot \bar{t}$.*

Proof. The proof is almost the same as Proposition D.1. Let $\bar{c} = (c_1, \ldots, c_m)$, $\bar{d} = (d_1, \ldots, d_n)$ and $\bar{t} = ({\blacktriangleright}t_1, \ldots, {\blacktriangleright}t_k)$. Let $\bar{M} \in \mathbb{N}^m$ and $\bar{N} \in \mathbb{N}^n$ and $\bar{K} \in \mathbb{N}^k$ such that

$$\begin{aligned} \bar{M} \cdot \bar{1} + \bar{N} \cdot \bar{1} + \bar{K} \cdot \bar{1} &\geq 2(\bar{c} \cdot \bar{1})(\bar{d} \cdot \bar{1} + \bar{t} \cdot \bar{1}) + 3, \\ \bar{K} \cdot \bar{1} &\geq \bar{c} \cdot \bar{1} \end{aligned}$$

The "only if" part is straightforward. If there exists a $(\bar{c}, (\bar{d}, \bar{t}))$-biregular graph of size $(\bar{M}, (\bar{N}, \bar{K}))$, then the number of edges in the graph is $\bar{M} \cdot \bar{c}$, which should be $\geq (\bar{N} \cdot \bar{d}) + (\bar{K} \cdot \bar{t})$

For the "if" direction, let $\bar{M} = (M_1, \ldots, M_m)$, $\bar{N} = (N_1, \ldots, N_n)$, $\bar{K} = (K_1, \ldots, K_k)$ such that $\bar{M} \cdot \bar{c} \geq (\bar{N} \cdot \bar{d}) + (\bar{K} \cdot \bar{t})$. We are going to construct a $(\bar{c}, (\bar{d}, \bar{t}))$-biregular graph of size $(\bar{M}, (\bar{N}, \bar{K}))$. First, we construct the following graph, as illustrated below. On the left side there are $\bar{M} \cdot \bar{c}$ vertices, each of degree 1. On the right side there are N_1 vertices of degree d_1, N_2 vertices of degree d_2, and so on; and K_1 vertices of degree t_1, K_2 vertices of degree t_2, and so on; and $\bar{M} \cdot \bar{c} - (\bar{N} \cdot \bar{d} + \bar{K} \cdot \bar{t})$ vertices of degree 1. We call the edges adjacent to those $\bar{M} \cdot \bar{c} - (\bar{N} \cdot \bar{d} + \bar{K} \cdot \bar{t})$ vertices the *auxiliary* edges, and the vertices the *auxiliary* vertices.



Then we partition the vertices on the left into $V_1 \cup \cdots \cup V_m$ as follows. The first group V_1 consists of $M_1 c_1$ vertices, the second group V_2 consists of $M_2 c_2$ vertices, and so on. Then we merge every c_1 vertices in V_1 into one vertex, every c_2 vertices in V_2 into one vertex, and so on. The merging is done like in the proof of Lemma D.2. After such merging, on the left side

there are M_1 vertices of degree c_1, M_2 vertices of degree c_2 and so on.

However, with such merging, we can have parallel edges between two vertices. If so, we apply the standard trick used in Lemma D.2. Suppose we have parallel edges between the vertices v and u. We pick a vertex u' of non zero degree with distance ≥ 2 from either both u or v. Such a vertex exists since the number of vertices is $\bar{M}\cdot\bar{1}+\bar{N}\cdot\bar{1}+\bar{K}\cdot\bar{1} \geq 2(\bar{c}\cdot\bar{1})(\bar{d}\cdot\bar{1}+\bar{t}\cdot\bar{1})+3$, and the number of vertices of distance 2 from both u and v is $\leq 2(\bar{c}\cdot\bar{1})(\bar{d}\cdot\bar{1}+\bar{t}\cdot\bar{1})+2$.

Pick an edge (u', v') for some neighbour v' of u'. We replace the edges (v, u) and (v', u') with (v, u') and (v', u). We do this until there is no more parallel edges.

Hence we get a bipartite graph $G = (U, V \cup W \cup W_0, E)$ such that

- there is a partition $U = U_1 \cup \cdots \cup U_m$, $V = V_1 \cup \cdots \cup V_n$ and $W = W_1 \cup \cdots \cup W_k$;
- for every $i \in \{1, \ldots, m\}$, for every vertex $u \in U_i$, $\deg(u) = c_i$;
- for every $i \in \{1, \ldots, n\}$, for every vertex $v \in V_i$, $\deg(v) = d_i$;
- for every $i \in \{1, \ldots, k\}$, for every vertex $w \in W_i$, $\deg(u) = t_i$;
- for every vertex $w \in W_0$, $\deg(w) = 1$.
  (The vertices in W_0 are the auxiliary vertices.)

Now we want to eliminate the auxiliary edges. Let (u, w') be an auxiliary edge, where $u \in U$ and $w' \in W_0$. Then we pick a vertex $w \in W$ such that $(u, w) \notin E$. Such a vertex w exists since, $\bar{K}\cdot\bar{1} \geq \bar{c}\cdot\bar{1} > \deg(u)$. Then delete (u, w') and add in the edge (v, w). We do this until there is no more auxiliary edges.

After such elimination of the auxiliary edges, we have a bipartite graph that

- on the left side there are M_i vertices of degree c_i, for each $i = 1, \ldots, m$;
- on the right side among the vertices u's there are N_i vertices of degree d_i, for each $i = 1, \ldots, n$;
- on the right side among the vertices w's there are K_i vertices of degree $\geq t_i$, for each $i = 1, \ldots, k$.

This is a $(\bar{c}, (\bar{d}, \bar{t}))$-biregular graph of size $(\bar{M}, (\bar{N}, \bar{K}))$. This completes our proof of Lemma D.7. ■

Lemma D.8 *Let $\bar{c} \in \mathbb{N}^m$, $\bar{s} \in ({}^{\blacktriangleright}\mathbb{N})^k$ and $\bar{d} \in \mathbb{N}^n$, $\bar{t} \in ({}^{\blacktriangleright}\mathbb{N})^l$. For every $\bar{M} \in \mathbb{N}^m$, $\bar{K} \in \mathbb{N}^k$, $\bar{N} \in \mathbb{N}^n$ and $\bar{L} \in \mathbb{N}^l$ such that*

$$\bar{M}\cdot\bar{1}+\bar{K}\cdot\bar{1}+\bar{N}\cdot\bar{1}+\bar{L}\cdot\bar{1} \quad \geq \quad 2(\bar{c}\cdot\bar{1}+\bar{s}\cdot\bar{1})(\bar{d}\cdot\bar{1}+\bar{t}\cdot\bar{1})+3,$$

the following holds.

(a) If $\bar{M}\cdot\bar{c}+\bar{K}\cdot\bar{s} > \bar{N}\cdot\bar{d}+\bar{L}\cdot\bar{t}$ and $\bar{L}\cdot\bar{1} \geq \bar{c}\cdot\bar{1}+\bar{s}\cdot\bar{1}$, then there exists a $((\bar{c}, \lfloor\bar{s}\rfloor), (\bar{d}, \bar{t}))$-biregular graph of size $((\bar{M}, \bar{K}), (\bar{N}, \bar{L}))$.

(b) If $\bar{M}\cdot\bar{c}+\bar{K}\cdot\bar{s} = \bar{N}\cdot\bar{d}+\bar{L}\cdot\bar{t}$, then there exists a $((\bar{c}, \lfloor\bar{s}\rfloor), (\bar{d}, \lfloor\bar{t}\rfloor))$-biregular graphof size $((\bar{M}, \bar{K}), (\bar{N}, \bar{L}))$.

(c) If $\bar{M}\cdot\bar{c}+\bar{K}\cdot\bar{s} < \bar{N}\cdot\bar{d}+\bar{L}\cdot\bar{t}$ and $\bar{K}\cdot\bar{1} \geq \bar{d}\cdot\bar{1}+\bar{t}\cdot\bar{1}$, then there exists a $((\bar{c}, \bar{s}), (\bar{d}, \lfloor\bar{t}\rfloor))$-biregular graphof size $((\bar{M}, \bar{K}), (\bar{N}, \bar{L}))$.

Proof. Parts (a) and (c) follow from Lemma D.7, while part (b) from Lemma D.2. ■

Proof of Theorem D.6. Without loss of generality, we are going to construct a formula $\mathsf{BiREG}_{(\bar{c},\bar{s}),(\bar{d},\bar{t})}(\bar{X},\bar{Y})$, where $\bar{c} \in \mathbb{N}^m$, $\bar{s} \in ({}^{\blacktriangleright}\mathbb{N})^k$ and $\bar{d} \in \mathbb{N}^n$, $\bar{t} \in ({}^{\blacktriangleright}\mathbb{N})^l$ and $\bar{X} = (X_1, \ldots, X_{m+k})$ and $\bar{Y} = (Y_1, \ldots, Y_{n+l})$. It should be noted that it is possible that any of m, k, n, l is zero.

We will also denote $\bar{X} = (\bar{X}_1, \bar{X}_2)$, where $\bar{X}_1 = (X_1, \ldots, X_m)$ is the vector of the first m variables from $\bar{X}$, while $\bar{X}_2 = (X_{m+1}, \ldots, X_{m+k})$ the vector of the last k variables from $\bar{X}$. Similarly, we also denote $\bar{Y} = (\bar{Y}_1, \bar{Y}_2)$, where $\bar{Y}_1 = (Y_1, \ldots, Y_n)$ is the vector of the first m variables from $\bar{Y}$, while $\bar{Y}_2 = (Y_{n+1}, \ldots, Y_{n+l})$ the vector of the last k variables from $\bar{Y}$.

Without loss of generality, we may make the following assumptions.

- Both $\bar{c}$ and $\bar{d}$ do not contain zero entry.
  If say $c_m = 0$, then we can add the inequality $X_m \geq 0$ and ignore that entry. Formally, we can define $\mathsf{BiREG}_{(\bar{c},\bar{s}),(\bar{d},\bar{t})}(\bar{X},\bar{Y})$ as follows.

$$X_m \geq 0 \quad \wedge \quad \mathsf{BiREG}_{(c_1,\ldots,c_{m-1},\bar{s}),(\bar{d},\bar{t})}(X_1, \ldots, X_{m-1}, X_{m+1}, \ldots, X_{m+k}, \bar{Y})$$

  We can do similar trick if $\bar{d}$ contains zero entry.

- Both $\bar{s}$ and $\bar{t}$ do not contain ${}^{\blacktriangleright}0$ entry.
  If say $s_k = {}^{\blacktriangleright}0$, then we define $\mathsf{BiREG}_{(\bar{c},\bar{s}),(\bar{d},\bar{t})}(\bar{X},\bar{Y})$ as follows.

$$\begin{aligned} \exists Z_1\, \exists Z_2 \qquad & Z_1 \geq 0 \quad \wedge \quad Z_1 + Z_2 = X_{m+k} \\ & \wedge \quad \mathsf{BiREG}_{(\bar{c},{}^{\blacktriangleright}s_1,\ldots,{}^{\blacktriangleright}s_{k-1},{}^{\blacktriangleright}1),(\bar{d},\bar{t})}(X_1, \ldots, X_{m+k-1}, Z_2, \bar{Y}) \end{aligned}$$

  Intuitively, the set of vertices with degree at least zero can be divided into two: one set of Z_1 vertices with degree zero and the other of Z_2 vertices with degree at least one. We can do similar trick if $\bar{t}$ contains ${}^{\blacktriangleright}0$ entry.

Lemmas D.7 and D.8 immediately gives us a Presburger formula that characterises the existence of $((\bar{c},\bar{s}),(\bar{d},\bar{t}))$-biregular graph of size $((\bar{M},\bar{K}),(\bar{N},\bar{L}))$ when

$$\begin{aligned} \bar{M}\cdot\bar{1} + \bar{K}\cdot\bar{1} + \bar{N}\cdot\bar{1} + \bar{L}\cdot\bar{1} &\geq 2(\bar{c}\cdot\bar{1} + \bar{s}\cdot\bar{1})(\bar{d}\cdot\bar{1} + \bar{t}\cdot\bar{1}) + 3 \\ \bar{K}\cdot\bar{1} &\geq \bar{d}\cdot\bar{1} + \bar{t}\cdot\bar{1} \\ \bar{L}\cdot\bar{1} &\geq \bar{c}\cdot\bar{1} + \bar{s}\cdot\bar{1} \end{aligned}$$

What is left is to handle one of the following cases.

(a) $\bar{M}\cdot\bar{1} + \bar{K}\cdot\bar{1} + \bar{N}\cdot\bar{1} + \bar{L}\cdot\bar{1} < 2(\bar{c}\cdot\bar{1} + \bar{s}\cdot\bar{1})(\bar{d}\cdot\bar{1} + \bar{t}\cdot\bar{1}) + 3$.

(b) Either $\bar{K}\cdot\bar{1} < \bar{d}\cdot\bar{1} + \bar{t}\cdot\bar{1}$ or $\bar{L}\cdot\bar{1} < \bar{c}\cdot\bar{1} + \bar{s}\cdot\bar{1}$.

Case (a) is trivial. We enumerate all those vectors $((\bar{M},\bar{K}),(\bar{N},\bar{L}))$ such that there exists a $(\bar{c},\bar{s}),(\bar{d},\bar{t})$-biregular graph of size $((\bar{M},\bar{K}),(\bar{N},\bar{L}))$ inside the formula $\mathsf{BiREG}_{(\bar{c},\bar{s}),(\bar{d},\bar{t})}(\bar{X},\bar{Y})$.

For case (b), we define the notion of partial graph for bipartite graph similar to the one defined in the proof of Theorem 4.2.

For completeness, we define it below. A *partial graph for bipartite graph* is a tuple (P, f, g, m, n), where

- $P = (U, V, E)$ is a bipartite graph;

- f is a function $U \times \{1, \ldots, n\} \to \mathbb{B}$;
- g is a function $V \times \{1, \ldots, m\} \to \mathbb{B}$.

A completion of the partial graph (P, f, g, m, n) is a bipartite graph $G = (U', V', E')$ such that

- P is a subgraph of G, i.e., $U \subseteq U'$, $V \subseteq V'$ and $E \subseteq E'$;
- there is a partition of the vertices $V' \setminus V = V_1 \cup \cdots \cup V_n$, where for each $j = 1, \ldots, n$, for each vertex $u \in U$ the number of vertices in V_j adjacent to u is $f(u, j)$;
- there is a partition of the vertices $U' \setminus U = U_1 \cup \cdots \cup U_m$, where for each $j = 1, \ldots, m$, for each vertex $v \in V$ the number of vertices in U_j adjacent to v is $g(v, j)$.

Using this notion of partial graph, in a manner similar to the proof of Theorem 4.2 we can similarly encode the part of $(\bar{c}, \bar{s}), (\bar{d}, \bar{t})$-biregular graph with a fixed size into the Presburger formula $\mathsf{BiREG}_{(\bar{c},\bar{d})}(\bar{X}, \bar{Y})$. The details are omitted.

Proof of Theorem 5.1. In the same manner as in the proof of Theorem D.4, Theorem D.6 can be generalised to obtain Theorem 5.1, the details of which are omitted.

E Proof of Theorem 5.2

The proof is very similar to the proof of Theorem 4.2, so we just sketch it here. The following lemma is the bipartite counterpart of Lemma C.1. The proof is also very similar to the proof of Lemma C.1.

Lemma E.1 *Let $G = (U, V, E_1, \ldots, E_\ell)$ be an ℓ-type (C, D)-biregular graph, where $C = \begin{pmatrix} \bar{c}_1 \\ \cdots \\ \bar{c}_\ell \end{pmatrix} \in \mathbb{B}^{\ell \times m}$, $D = \begin{pmatrix} \bar{d}_1 \\ \cdots \\ \bar{d}_\ell \end{pmatrix} \in \mathbb{B}^{\ell \times n}$ and $U = U_1 \cup \cdots \cup U_m$ and $V = V_1 \cup \cdots \cup V_n$ is the partition of the regularity. Suppose that for each i, j, we have*

$$|U_i|, \quad |V_j| \quad \geq \quad (\sum_i \bar{c}_i) \cdot \bar{1} + (\sum_j \bar{d}_j) \cdot \bar{1} + 1.$$

If G is a complete bipartite graph, then for every $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, there exists $l \in \{1, \ldots, \ell\}$ such that both $C_{l,i}, D_{l,j} \in {}^{\blacktriangleright}\mathbb{N}$.

Similarly as in the regular graph, we say that a pair of matrices $(C, D) \in \mathbb{B}^{\ell \times m} \times \mathbb{B}^{\ell \times n}$ is an *easy pair of matrices*, if for every $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ there exists $l \in \{1, \ldots, \ell\}$ such that both $C_{l,i}, D_{l,j} \in {}^{\blacktriangleright}\mathbb{N}$. The following lemma can be proved in the same way as Lemma C.2.

Lemma E.2 *Let (C, D) be an easy pair of matrices. Then the following holds. There exists a (C, D)-biregular-complete graph of size $(\bar{M}, \bar{N})$ if and only if $\mathsf{BiREG}_{C,D}(\bar{M}, \bar{N})$ holds.*

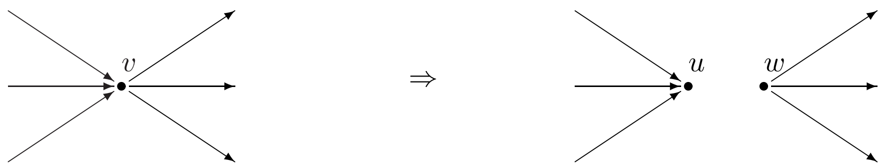
Now Theorem 5.2 can be easily proved in a similar manner as in Appendix C, with the difference that now we use partial graph for bipartite graph defined in Subsection D.4.

F Proof of Theorems 7.1 and 7.2

The proof is by observing that the existence of a (C, D)-directed-regular graph of size $\bar{N}$ is equivalent to the existence of a (C, D)-biregular graph of size $(\bar{N}, \bar{N})$. We explain it more precisely below.

- Suppose $G = (V, E_1, \ldots, E_\ell)$ is a (C, D)-directed-regular graph of size $\bar{N}$.

  Then, for every vertex $v \in V$, we "split" it into two vertices u and w such that u is only adjacent to the incoming edges of v and w to the outgoing edges of v. See the illustration below. The left-hand side shows the vertex v before the splitting, and the right-hand side shows the vertices u and w after the splitting.

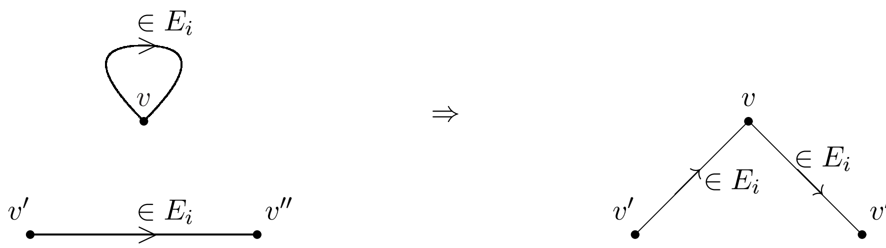
  

  Let U be the set of vertices u's and W the set of vertices w's after splitting all the vertices in V. Ignoring the orientation of the edges, the resulting graph is a bipartite graph with vertices $U \cup W$ and it is a (C, D)-biregular graph of size $(\bar{N}, \bar{N})$.

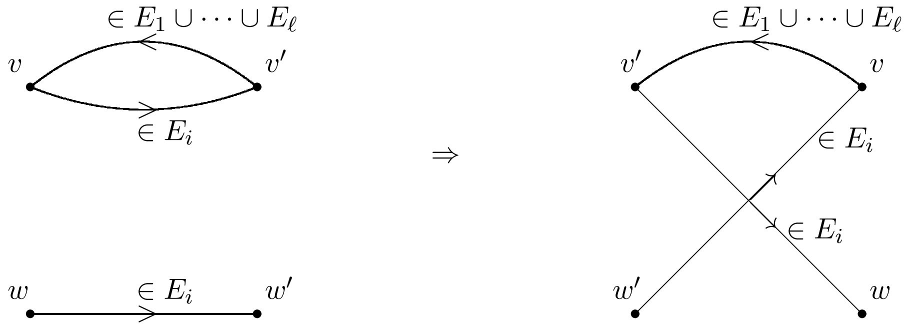
- Suppose $G = (U, W, E_1, \ldots, E_\ell)$ is a (C, D)-biregular graph of size $(\bar{N}, \bar{N})$. Let $U = U_1 \cup \cdots \cup U_m$ and $W = W_1 \cup \cdots \cup W_m$ be the partition of (C, D)-biregularity. We denote by $U_i = \{u_{i,1}, \ldots, u_{i,K_i}\}$ and $W_i = \{w_{i,1}, \ldots, w_{i,K_i}\}$ for each $i = 1, \ldots, m$.

  Now we put the orientation on all the edges from U to W. Then we merge every two vertices $u_{i,j}$ and $w_{i,j}$ into one vertex $v_{i,j}$. This way, we obtain a (C, D)-directed-regular graph $G = (V, E_1, \ldots, E_\ell)$ with $V = V_1 \cup \cdots \cup V_m$ be the partition of regularity where $V_i = \{v_{i,1}, \ldots, v_{i,K_i}\}$ for each $i = 1, \ldots, m$.

  However, with such merging it is possible that there is a self-loop (v, v) in G or a pair of edges $(v, v'), (v', v) \in E_1 \cup \cdots \cup E_\ell$. We can get rid of the self-loop (v, v) without violating the (C, D)-directed-regularity as follows. The trick is similar to the one used before. Assuming that the size of each $|V_1|, \ldots, |V_m|$ is big enough and there are enough edges in each $E_1, \ldots, E_\ell$, there is an edge (v', v'') of the same type. Deleting the edge (v, v) and (v', v''), and adding the edges (v', v) and (v, v''), we obtain a (C, D)-directed-regular graph with one less self-loop. We do this repeatedly until there is no more self-loop. See the illustration below.



Similarly, we can get rid of a pair of edges $(v, v'), (v', v) \in E_1 \cup \cdots \cup E_\ell$ without violating the (C, D)-directed-regularity as follows. Again, the trick is similar to the one used before. Assuming that the size of each $|V_1|, \ldots, |V_m|$ is big enough and there are enough edges in each $E_1, \ldots, E_\ell$, there is an edge (w, w') of the same type. Deleting the edge (v, v') and (w, w'), and adding the edges (v, w') and (w, v'), we obtain a (C, D)-directed-regular graph with one less parallel edges. We do this repeatedly until there is no more parallel edges. See the illustration below.



If some sets V_i or $|E_i|$ are of a fixed size, those can be encoded in a partial graph in the same manner discussed in Section C.

We omit the technical details since we essentially run through the same argument used in the previous sections.

G The details for the proof of Proposition 6.1

The formal proof is as follows. Recall that $\mathcal{P} = \{P_1, P_2, \ldots\}$ be set of unary predicates used in ϕ. Abusing the notation, we let PREB_ϕ itself to denote the set $\{N \mid \mathsf{PREB}_\phi(N) \text{ holds}\}$. We claim that $\mathsf{PREB}_\phi = \textsc{Spec}(\phi)$.

We first show the $\subseteq$ direction. Let $N \in \mathsf{PREB}_\phi$. Let $\bar{M} = (M_{T_1,f_1}, \ldots, M_{T_n,f_m})$ be the witnesses to $\bar{X}$ such that $\mathsf{PREB}_\phi(N)$ holds. In the following we are going to write $\bar{M}_T$ to denote $(M_{T,f_1}, \ldots, M_{T,f_m})$ for every type $T \in \mathcal{T}_\phi$.

By definition $N = \sum_{(T,f)} M_{(T,f)}$. We take a set V of N vertices and we partition $V = V_{(T_1,f_1)} \cup \cdots \cup V_{(T_n,f_m)}$ such that $|V_{(T,f)}| = M_{(T,f)}$ for each $T \in \mathcal{T}_\phi$ and $f \in \mathcal{F}$. We denote by $V_T = V_{(T,f_1)} \cup \cdots \cup V_{(T,f_m)}$ for each $T \in \mathcal{T}_\phi$.

Now since $\mathsf{CON}(\bar{M})$ holds, by Theorems 4.2, for each $T \in \mathcal{T}_\phi$, there exists a D_T-regular-complete graph $G_T = (V_T, R_{T,1}, \ldots, R_{T,\ell})$ of size $\bar{M}_T$, with $V_T = V_{(T,f_1)} \cup \cdots \cup V_{(T,f_m)}$ be the partition of D_T-regularity. Similarly, by Theorem 5.2 for each $S, T \in \mathcal{T}_\phi$, there exists a $(D_{S\to T}, D_{T\to S})$-biregular-complete graph $G_{S,T} = (V_S, V_T, R_{S,T,1}, \ldots, R_{S,T,\ell})$ of size

$(\bar{M}_S, \bar{M}_T)$, with $V_T = V_{(T,f_1)} \cup \cdots \cup V_{(T,f_m)}$ and $V_S = V_{(S,f_1)} \cup \cdots \cup V_{(S,f_m)}$ be the partition of $(D_{S\to T}, D_{T\to S})$-biregularity.

Let G be the overall graph $G = (V, R_1, \ldots, R_\ell)$, where

$$R_i \;=\; \bigcup_T R_{T,i} \;\cup\; \bigcup_{S,T} R_{S,T,i}$$

Then we label each vertex $v \in V$ with a subset of $\mathcal{S}$ as follows. For each $T \in \mathcal{T}_\phi$, for each $v \in V_T$, we "declare" that v is labeled with a unary predicate $P \in \mathcal{P}$ if and only if $P \in T$.

We claim that $G \models \phi$. For that it is sufficient to show that for each $T \in \mathcal{T}_\phi$, for each $v \in V_T$, $\mathsf{type}_G(v) = T$. The proof is divided into three cases.

- For each unary predicate $P \in \mathcal{P}$, it is by our labelling of the vertices of G that $P(v)$ holds in G if and only if $P \in T$.
- For each $\Diamond_R^l\ \mu \in T$, we have $\sum_{T' \text{ s.t. } T' \ni \mu}\ f(T, R, T') \geq l$ number of R-edges adjacent to v. Since every function $f \in \mathcal{F}$ is consistent, $\Diamond_R^l\ \mu \in \text{type}(v)$.
- Similary, for each $\Diamond_R^l\ \mu \notin T$, and hence $\neg(\Diamond_R^l\ \mu) \in T$, we have $\sum_{T' \text{ s.t. } T' \ni \mu}\ f(T, R, T') \leq l - 1$, number of R-edges adjacent to v. Since every function $f \in \mathcal{F}$ is consistent, $\neg\Diamond_R^l\ \mu \in \text{type}(v)$.

Therefore the graph $G \models \phi$, and hence $n \in \textsc{Spec}(\phi)$.

Now we prove the direction $\supseteq$. Suppose $\mathfrak{A} \models \phi$ is of size N. Let $\bar{M} = (M_{(T_1,f_1)}, \ldots, M_{(T_n,f_m,)})$ where $M_{(T,f)}$ be the number of elements of type T in which there exists $f(T, R, S)$ number of R-edges going to elements of type S. Take each $M_{(T,f)}$ to be the witness for $X_{(T,f)}$ for each $T \in \mathcal{T}_\phi$ and $f \in \mathcal{F}$. By Theorems 4.2 and 5.2, $\mathsf{CON}(\bar{M})$ holds. Moreover, since $\mathfrak{A} \models \phi$, then $\mathsf{PREB\text{-}Atom}_\phi(\bar{M})$ holds. This completes the proof of our proposition.

H The details for the proof of Proposition 8.1

The formal proof is as follows. Recall that $\mathcal{P} = \{P_1, P_2, \ldots\}$ be set of unary predicates used in ϕ. Abusing the notation, we let PREB_ϕ itself to denote the set $\{n \mid \mathsf{PREB}_\phi(n) \text{ holds}\}$. We claim that $\mathsf{PREB}_\phi = \textsc{Spec}(\phi)$.

We first show the $\subseteq$ direction. Let $N \in \mathsf{PREB}_\phi$. Let $\bar{M} = (M_{T_1,f_1}, \ldots, M_{T_n,f_m})$ be the witnesses to $\bar{X}$ such that $\mathsf{PREB}_\phi(N)$ holds. In the following we are going to write $\bar{M}_T$ to denote $(M_{T,f_1}, \ldots, M_{T,f_m})$ for every type $T \in \mathcal{T}_\phi$.

By definition $N = \sum_{(T,f)} M_{(T,f)}$. We take a set V of N vertices and we partition $V = V_{(T_1,f_1)} \cup \cdots \cup V_{(T_n,f_m)}$ such that $|V_{(T,f)}| = M_{(T,f)}$ for each $T \in \mathcal{T}_\phi$ and $f \in \mathcal{F}$. We denote by $V_T = V_{(T,f_1)} \cup \cdots \cup V_{(T,f_m)}$ for each $T \in \mathcal{T}_\phi$.

Now since $\mathsf{CON}(\bar{M})$ holds, by Theorems 7.2, for each $T \in \mathcal{T}_\phi$, there exists a $(D_T, \overleftarrow{D}_T)$-directed-regular-complete graph $G_T = (V_T, R_{T,1}, \ldots, R_{T,\ell})$ of size $\bar{M}_T$, with $V_T = V_{(T,f_1)} \cup \cdots \cup V_{(T,f_m)}$ be the partition of $(D_T, \overleftarrow{D}_T)$-regularity. This means that for every vertex $v \in V_{T,f_i}$, for every $R \in \{R_1, \ldots, R_\ell\}$,

- out-$\deg_R(v)$ in the graph G_T is $f_i(T, R, T)$;
- in-$\deg_R(v)$ in the graph G_T is $f_i(T, \overleftarrow{R}, T)$.

Now let $\tilde{G}_T = (V_T, R_{T,1}, \ldots, R_{T,\ell}, \overleftarrow{R}_{T,1}, \ldots, \overleftarrow{R}_{T,\ell})$ be the graph obtained by taking $\overleftarrow{R}_i$ as the reverse of R_i. Hence, we have for each vertex $v \in V_T$,

- $\text{out-deg}_R(v) = \text{in-deg}_{\overleftarrow{R}}(v)$ in the graph $\tilde{G}_T$;
- $\text{in-deg}_R(v) = \text{out-deg}_{\overleftarrow{R}}(v)$ in the graph $\tilde{G}_T$.

Similarly, by Theorem 5.2 for each $T_i, T_j \in \mathcal{T}_\phi$, where $j \le i-1$, there exists a $(D_{T_i \to T_j}, \overleftarrow{D}_{T_i \to T_j})$-biregular-complete graph

$$G_{T_i,T_j} = (V_{T_i}, V_{T_j}, R_{T_i,T_j,1}, \ldots, R_{T_i,T_j,\ell}, \overleftarrow{R}_{T_i,T_j,1}, \ldots, \overleftarrow{R}_{T_i,T_j,\ell})$$

of size $(\bar{M}_S, \bar{M}_T)$, with $V_{T_i} = V_{(T_i,f_1)} \cup \cdots \cup V_{(T_i,f_m)}$ and $V_{T_j} = V_{(T_j,f_1)} \cup \cdots \cup V_{(T_j,f_m)}$ be the partition of $(D_{T_i \to T_j}, \overleftarrow{D}_{T_i \to T_j})$-biregularity. This means that for every $R \in \{R_1, \ldots, R_\ell, \overleftarrow{R}_1, \ldots, \overleftarrow{R}_\ell\}$,

- for every vertex $v \in V_{T_i,f}$, $\text{out-deg}_R(v)$ in the graph G_{T_i,T_j} is $f(T_i, R, T_j)$;
- for every vertex $v \in V_{T_j,f}$, $\text{out-deg}_R(v)$ in the graph G_{T_i,T_j} is $f(T_i, \overleftarrow{R}, T_j)$.

Then, we put orientation in all the edges in the graph G_{T_i,T_j} going from V_{T_i} to V_{T_j}. And now let $\tilde{G}_{T_i,T_j}$ be the graph adding the reverse edges from V_{T_i} to V_{T_j} by putting (u,v) into $\overleftarrow{R}$ in the graph $\tilde{G}_{T_i,T_j}$ whenever $(v,u) \in R(G_{T_i,T_j})$ for each $R \in \mathcal{R}$.

Hence, we have for each vertex $v \in V_{T_i} \cup V_{T_j}$, for each $R \in \mathcal{R}$

- $\text{out-deg}_R(v) = \text{in-deg}_{\overleftarrow{R}}(v)$ in the graph $\tilde{G}_{T_i,T_j}$;
- $\text{in-deg}_R(v) = \text{out-deg}_{\overleftarrow{R}}(v)$ in the graph $\tilde{G}_{T_i,T_j}$.

Let G be the combination of all the graphs $G = (V, R_1, \ldots, R_\ell, \overleftarrow{R}_1, \ldots, \overleftarrow{R}_\ell)$, where for each $R \in \mathcal{R}$,

$$R \;=\; \bigcup_T R(\tilde{G}_T) \;\cup\; \bigcup_{T_i,T_j} R(\tilde{G}_{T_i,T_j})$$

Moreover, we also label each vertex $v \in V$ with a subset of $\mathcal{S}$ as follows. For each $T \in \mathcal{T}_\phi$, for each $v \in V_T$, we "declare" that v is labeled with a unary predicate $P \in \mathcal{P}$ if and only if $P \in \mathcal{P}$.

We claim that $G \models \phi$. For that it is sufficient to show that for each $T \in \mathcal{T}_\phi$, for each $v \in V_T$, $\mathsf{type}_G(v) = T$. The proof is divided into three cases.

- For each unary predicate $P \in \mathcal{P}$, it is by our labelling of the vertices of G that $P(v)$ holds in G if and only if $P \in T$.
- For each $\Diamond_R^l\, \mu \in T$, where $R \in \mathcal{R}$ we have

  $$\sum_{T' \text{ s.t. } T' \ni \mu} f(T, R, T') \;\;\ge\;\; l$$

  number of outgoing R-edges from v. Since every function $f \in \mathcal{F}$ is consistent, $\Diamond_R^l\, \mu \in \mathsf{type}(v)$.
- Similary, for each $\Diamond_R^l\, \mu \notin T$, and hence $\neg(\Diamond_R^l\, \mu) \in T$, where $R \in \mathcal{R}$ we have

  $$\sum_{T' \text{ s.t. } T' \ni \mu} f(T, R, T') \;\;\ge\;\; l$$

  number of outgoing R-edges from v. Since every function $f \in \mathcal{F}$ is consistent, $\Diamond_R^l\, \mu \in \mathsf{type}(v)$.

Therefore the graph $G \models \phi$, and hence $N \in \textsc{Spec}(\phi)$.

Now we prove the direction $\supseteq$. Suppose $\mathfrak{A} \models \phi$ is of size N. Let $\bar{M} = (M_{(T_1,f_1)}, \ldots, M_{(T_n,f_m,)})$ where $M_{(T,f)}$ be the number of elements of type T from which there exist $f(T, R, S)$ number of outgoing R-edges towards the elements of type S. Take each $M_{(T,f)}$ to be the witness for $X_{(T,f)}$ for each $T \in \mathcal{T}_\phi$ and $f \in \mathcal{F}$. It immediately follows from Theorems 7.2 and 5.2 that $\mathsf{CON}(N, \bar{M})$ holds. Moreover, since $\mathfrak{A} \models \phi$, then $\mathsf{PREB\text{-}Atom}_\phi(\bar{M})$ holds. This completes the proof of our proposition.